Investigations into Selective Metabolic Aspects of Bifidobacteria: Carbohydrate Metabolism, Fatty Acid Biosynthesis and Plasmid Biology

Ollscoil na hÉireann, Corcaigh

THE NATIONAL UNIVERSITY OF IRELAND, CORK

A Thesis presented to the National University of Ireland for the Degree of Doctor of Philosophy by

Kerry Joan O' Connell, B.Sc

Department of Microbiology

University College Cork

Supervisors: Prof. Douwe van Sinderen and Prof. Catherine Stanton

January 2014

Declaration

I hereby declare that the research presented in this thesis is my own work and effort and that it has not been submitted for any other degree, either at University College Cork or elsewhere. Wherever contributions of others are involved, every effort is made to indicate this clearly, with due reference to the literature and acknowledgement of collaborative research and discussions.

This work was completed under the guidance of Prof. Douwe van Sinderen and Prof. Catherine Stanton at the Alimentary Pharmabiotic Centre, Biosciences Institute (Microbiology Department), University College Cork.

Signature:……

Date:……

LIST OF FIGURES

LIST OF TABLES

This thesis is dedicated to my parents Michael and Siobhan O' Connell

"For your endless love and support"

The Road Not Taken

Two roads diverged in a yellow wood,

And sorry I could not travel both

And be one traveler, long I stood

And looked down one as far as I could

To where it bent in the undergrowth;

Then took the other, as just as fair,

And having perhaps the better claim

Because it was grassy and wanted wear,

Though as for that the passing there

Had worn them really about the same,

And both that morning equally lay

In leaves no step had trodden black.

Oh, I kept the first for another day!

Yet knowing how way leads on to way

I doubted if I should ever come back.

I shall be telling this with a sigh

Somewhere ages and ages hence:

Two roads diverged in a wood, and I,

I took the one less traveled by,

And that has made all the difference.

-Robert Frost

LIST OF PUBLICATIONS

O'Connell, K. J., Motherway, M. O., Hennessey, A. A., Brodhun, F., Ross, R. P., Feussner, I., Stanton, C., Fitzgerald, G. F. & van Sinderen, D. (2013). Identification and Characterisation of an Oleate Hydratase-encoding Gene From *Bifidobacterium breve*. *Bioengineered* **4**, 313–321.

O'Connell, K. J., O'Connell Motherway, M., O'Callaghan, J., Fitzgerald, G. F., Ross, R. P., Ventura, M., Stanton, C. & van Sinderen, D. (2013). Metabolism of Four α-glycosidic linkage-containing Oligosaccharides by *Bifidobacterium breve* UCC2003. *Applied and Environmental Microbiology* **79**, 6280-6292.

ABBREVIATIONS

ABC-type transporters: ATP-Binding Cassette Transporters

AckA: Acetate Kinase

Adh2: Aldehyde-Alcohol Dehydrogenase 2

Aga: α-Galactosidase;

Agl: α-Glucosidase

ALD: Adrenoleukodystrophy

APC: Alimentary Pharmabiotic Centre

ATP: Adenosine Triphosphate

Bgl: β-Glucosidase

BLAST: Basic Local Alignment Search Tool

c9, t11: *cis*-9 *trans*-11

CCR: Carbon Catabolite Repression

CFU: Colony Forming Units

CLA: Conjugated Linoleic Acid

DHA: Docosahexaenoic Acid

DNA: Deoxyribonucleic Acid

DTT: Dithiothreitol

EDTA: Ethylenediaminetetraacetic acid

EMSA: Electrophoretic Mobility Shift Assay

EPA: Eicosapentaenoic Acid

FAD: Flavin Adenine Dinucleotide

FOS: Fructooligosaccharides

F6PPK: Fructose-6-Phosphoketolase

Frk: Fruktokinase

FucA: L-Fuculose-1-P-Aldolase

FucI: L-Fucose Isomerase

FucK: L-Fuculose Kinase

FucO: Lactaldehyde Reductase

GalE1: UDP-glucose 4-epimerase

GalK: Galactokinase;

GalM: Galactose Mutarotase;

GAPDH: Glyceraldehyde-3-Phosphate Dehydrogenase C

GHs: Glycosyl Hydrolases

GIT: Gastrointestinal tract

GLC: Gas Liquid Chromatography

GlkA: Glucokinase

Gnt: 6-Phosphogluconate Dehydrogenase

GOS: Galactooligosaccharides

Gpi: Glucose 6-Phosphate isomerase;

HMOs: Human Milk Oligosaccharides

HPAEC-PAD: High Performance Anion Exchange Chromatography with Pulsed Amperometric Detection

10-HSA: 10-Hydroxystearic acid

HTH: Helix Turn Helix

IPTG: Isopropyl-β-D-thiogalactopyranoside

JCM: Japan Collection of Microorganisms

LA: Linoleic acid

LAI: Linoleic acid isomerase

LB: Luria Bertani broth

Ldh2: Lactate dehydrogenase

LMG: Belgian coordinated Collection of Microorganisms

LNBP: Lacto-N-biose phosphorylase

MCRA: Myosin Cross Reactive Antigen

MRS: Man Rogosa and Sharpe medium

mMRS: Modified Man Rogosa and Sharpe medium

MUFA: Monounsaturated fatty acids

MW: Molecular weight

NADH: Nicotinamide adenine dinucleotide

NADPH: Nicotinamide adenine dinucleotide phosphate

NCBI: National Centre for Biotechnology Information

NCFB: National Collection of Food Bacteria

NCIMB: National Collection of Industrial and Marine Bacteria

NCTC: National Collection of Type Cultures

NIZO: Nizo Food Research

OA: Oleic acid

OD: Optical density

ORF: Open reading frame

OriC: Origin of chromosomal replication

PCR: Polymerase Chain Reaction

PEP-PTS: Phosphoenolpyruvate-dependent phosphotransferase systems

Pfl: Formate acetyltransferase

Pgk: Phosphoglyceric kinase

Pgm: Phosphoglucomutase

Pi: Phosphate

Poly[d(I-C)]: Poly-desoxy-inosyl-desoxy-cytidylacid, Sodium salt

PRL: Culture collection of probiogenomics University of Parma.

PUFA: Polyunsaturated fatty acid

RBS: Ribosomal binding site

RCA: Reinforced clostridial agar

RCM: Reinforced clostridial media

Rk: Ribokinase

RNA: Ribonucleic acid

rRNA: Ribosomal RNA

ROK: Repressor open reading frame kinase

R5PI: Ribose-5-phosphate isomerase

R5PE: Ribulose-5-phosphate epimerase

SDS-PAGE: Sodium Dodecyl Sulphate Polyacrylamide Gel Electrophoresis

Tal: Transaldolase

t10, c12: *trans*-10, *cis*-12

Tet: Tetracycline

Tkt: Transketolase

TpiA: Triosephosphate isomerase

Tris-HCL: TRIS hydrochloride

t9, t11: *trans*-9, *trans*-11

UCC: University College Cork Culture Collection.

UgpA: UTP-glucose-1-phosphate uridylyltransferase

X-gal: 5-bromo-4-chloro-3-indolyl-D-galactopyranoside

XPPKT: Xylulose-5-phosphate/fructose-6-phosphate phosphoketolase

XylA: Xylose isomerase

XylB: Xylulose kinase

Zwf2: Glucose-6-phosphate 1-dehydrogenase

GENERAL ABSTRACT

The gastrointestinal tract (GIT) is a diverse ecosystem, and is colonised by a diverse array of bacteria, of which bifidobacteria are a significant component. Bifidobacteria are Gram-positive, saccharolytic, non-motile, non-sporulating, anaerobic, Y-shaped bacteria, which possess a high GC genome content. Certain bifidobacteria possess the ability to produce conjugated linoleic acid (CLA) from linoleic acid (LA) by a biochemical pathway that is hypothesised to be achieved via a linoleic isomerase. CLA has a number of published beneficial effects on the host, for example modification of body composition by supplementing CLA to the diet or inhibiting cancer by blocking growth and metastatic spread of tumours.

In Chapter two of this thesis we describe the CLA-producing capabilities of various *B. breve* strains, demonstrating that *B. breve* NCFB 2258 exhibits the highest level of CLA production. Since the gene predicted to encode a so-called myosin cross reactive antigen (*MCRA*) in *B. breve* NCFB 2258 shows high sequence homology to the linoleic acid isomerase of *Lactobacillus acidophilus* and *Lactobacillus reuteri* PYR8, an insertion mutation in the gene encoding MCRA was constructed to investigate its effect on CLA production. However, it was found that the *MCRA*-specifying gene is not involved in CLA production in *B. breve* NCFB 2258, and that this gene specifies an oleate hydratase involved in the conversion of oleic acid into 10-hydroxystearic acid.

Prebiotics are defined as non-digestible food ingredients that beneficially affect the host by selectively stimulating growth and/or activity of one or a limited number of bacteria in the colon. Many prebiotics represent oligo- or polysaccharides, which specifically target bifidobacteria due to their perceived probiotic effects on this

bacterial group. There is significant scientific and commercial interest in the identification of carbohydrates that can be developed as novel prebiotics with proven probiotic efficacy and bacterial target specificity. Key to the development of such novel prebiotics is to understand which carbohydrates support growth of bifidobacteria and how such carbohydrates are metabolised.

In Chapter 3 of this thesis we describe the identification and characterisation of two neighbouring gene clusters involved in the metabolism of raffinose-containing carbohydrates (plus related carbohydrate melibiose) and melezitose by *Bifidobacterium breve* UCC2003, which was used as our model strain and is an isolate from a nursling stool. The *rafABCD* gene cluster, whose transcription is specifically induced by growth on raffinose-containing carbohydrates and the related carbohydrate melibiose, is predicted to specify a GH family 27 α-galactosidase (*rafA*), a solute binding protein (*rafB*) and two permease proteins (*rafC* and *rafD*). The *melABCDE* gene cluster, transcription of which is specifically induced during growth on melezitose, is predicted to specify a solute binding protein (MelA) and two permease proteins (MelB and MelC), an α-glucosidase (MelD) belonging to GH family 13, and an α-galactosidase/raffinose synthase (MelE) of GH family 36.

The purified RafA protein was shown to hydrolyse stachyose and raffinose to produce sucrose and galactose, while also cleaving melibiose to its monosaccharide constituents glucose and galactose. Purified recombinant MelD protein was shown to fully hydrolyse melezitose into glucose and fructose, thereby demonstrating that this protein has both α-(1→2) and α-(1→3) glucosyl hydrolase activities. The putative α-galactosidase MelE failed to exhibit hydrolytic activity towards raffinose, stachyose or melibiose, however, MelE as well as RafA were shown to exhibit hydrolytic activity towards synthetic α-(1→4)- and α-(1→3)-galactobiose. Kinetic studies

identified that the preferred bond cleaved by MelD is the α-(1→2)-glucosidic linkage present in sucrose. In the case of RafA the preferred substrate was melibiose, followed by raffinose.

The fourth chapter of this thesis describes the analysis of transcriptional regulation of the *raf* and *mel* clusters. From our obtained information, it was concluded that the raffinose utilisation system is positively regulated by an activator protein, designated RafR, while the gene cluster associated with melezitose metabolism is negatively regulated by a LacI-type transcriptional regulator, designated MelR1. A second LacI-type transcriptional regulator, MelR2, encoded by a gene located immediately adjacent the MelR1-encoding gene, was shown to negatively regulate transcription of two genes that are positioned between the *melABCDE* gene cluster and the MelR2R1-encoding genes. Through a combination of *in silico* analysis, DNA/protein interaction and primer extension studies we determined that the MelR1 and MelR2 operator sequences partially overlap with the individual promoter sequences they regulate. Similar analyses identified the RafR binding operator sequence as being positioned just upstream of the *raf*B promoter. Furthermore, *B. breve* UCC2003 was shown to preferentially utilise melezitose over raffinose, indicating that this bifidobacterial strain operates an as yet unidentified carbon catabolite control mechanism.

In the final experimental chapter two putative *rep* genes, designated $repA_{7017}$ and $repB_{7017}$, are identified on the megaplasmid pBb7017 of *B. breve* JCM 7017, the first bifidobacterial megaplasmid to be reported. One of these, $repA_{7017}$, was subjected to an in-depth characterisation. *In silico* analysis, DNA/protein interaction and primer extension identified the transcriptional start site and binding region of the RepA protein and led to the hypothesis that this plasmid functions as a theta-type

replicating plasmid. The work described in this thesis has resulted in an improved understanding of bifidobacterial fatty acid and carbohydrate metabolism, and in the latter case also provided detailed insights into the carbohydrate-dependent regulation of different gene clusters involved in sugar metabolism. Furthermore, attempts were made to develop novel genetic tools, based on the pBb7017 replication functions which would facilitate the stable cloning of large DNA fragments, and which would represent a significant contribution towards the genetic manipulation of bifidobacteria and advance the functional investigation of genomic probiotic mechanisms of action.

CHAPTER I

General Introduction

1.1 INTRODUCTION

The gastrointestinal tract (GIT) represents an ecosystem that is colonised by a diverse array of bacteria, termed the gut microbiota, which in humans is estimated to outnumber the total counts of germ and somatic cells by approximately tenfold, and which encompasses approximately 100-fold more genes than the number present in the host genome (Bäckhed *et al.,* 2005). Prior to birth the foetus' gastrointestinal tract is considered to be essentially sterile, and the development of the gut microbiota commences immediately following birth as a result of contact with elements of the mother's vaginal and faecal microbiota during passage through the birth canal, as well as with other environmental microbes (Koenig *et al.,* 2011; Vaishampayan *et al.*, 2010).

The bacterial populations in the GIT can be divided into autochthonous and allochthonous bacteria, representing indigenous and transient microbes, respectively (Tannock, 2001), of which the former are presumed to be stably colonised in the human host (Zoetendal *et al.,* 1998). Determination of the faecal microbiota composition of healthy human adults has shown that the majority of microorganisms present are members of two distinct phyla, i.e. *Firmicutes* and *Bacteroides* (Eckburg *et al.,* 2005; Hold *et al.,* 2002). The faecal microbiota is presumed to reflect that of the gut itself and possesses a range of metabolic abilities that are otherwise absent in the host, and that are therefore considered indispensable for a healthy human existence (Savage, 2001).

The infant gut microbiota is thought to be sustained by specific oligo-saccharides present in mother's milk, and it is during this lactation period that bifidobacteria are the most abundant bacterial component of the gut microbiota (Gueimonde *et al.,*

2007). Several recent studies on the infant microbiota further reinforced the fact that bifidobacteria are predominant in the infant gut as determined by the analysis of faecal samples (Ishizeki *et al.*, 2013; Ruiz-Moyano *et al.*, 2013; Turroni *et al.*, 2012a). The prevalence of bifidobacteria in the infant gut has been linked to development and maturation of the immune system, and maintenance of proper gut function, although the precise mechanism(s) underlying such activities remain(s) to be elucidated (Klaenhammer *et al.*, 2012; Blum & Schiffrin, 2003; Ouwehand, 2007; Salminen *et al.*, 2005).

Heterotrophic bacteria such as bifidobacteria obtain energy from (limited) oxidation of organic compounds, and in particular carbohydrates. Oxidation of these compounds allows the synthesis of adenosine triphosphate (ATP) as the chemical energy source. This literature review will focus on carbohydrate and fatty acid metabolism in bifidobacteria.

1.2 GENERAL FEATURES OF BIFIDOBACTERIA

Bifidobacteria are Gram-positive, saccharolytic, non-motile, non-sporulating, anaerobic bacteria, which exhibit a bifurcated morphology, and possess a high G+C genome content (Ventura *et al.*, 2007b). Members of the genus *Bifidobacterium* are part of the phylum *Actinobacteria*, which is one of the largest phyla in the domain *Bacteria* (Ventura *et al.*, 2007a). Within the *Bifidobacterium* genus there are currently 48 species, many of which fall within six phylogenetic clusters, namely the *Bifidobacterium pseudolongum, Bifidobacterium pullorum, Bifidobacterium longum, Bifidobacterium adolenscentis, Bifidobacterium asteroides* and *Bifidobacterium boum* groups (Ventura *et al*., 2006). Bifidobacteria are present in different ecological niches: animal intestine (human, bovine, rabbit, murine, chicken and insect), oral

cavity, sewage, blood and food (Turroni *et al.*, 2008; Turroni *et al.*, 2012b; Ventura *et al.*, 2007a). They were first isolated over hundred years ago (Tissier, 1900) from the faeces of a breast-fed infant and were then named "*Bacillus bifidus*". Various bifidobacterial strains have been associated with a wide range of health-promoting activities, such as strengthening of the intestinal barrier, modulation of the immune system and exclusion of pathogens (Fanning *et al.*, 2012; Marco *et al.*, 2006; O'Hara & Shanahan, 2007; Turroni *et al.*, 2008; Ventura *et al.*, 2012; Ventura *et al.*, 2014).

A substantial amount of physiological and molecular biological information is available on *Bifidobacterium breve* UCC2003, an isolate from nursling stool (O'Connell Motherway *et al.*, 2011b), which has been demonstrated to metabolise various mono-, di-, oligo- and polysaccharides (Maze *et al.*, 2007; O'Connell *et al.*, 2013b; O'Connell Motherway *et al.*, 2008; O'Connell Motherway *et al.*, 2011a; Pokusaeva *et al.*, 2010; Pokusaeva *et al.*, 2011b; Ryan *et al.*, 2005).

1.2.1 BIFIDOBACTERIAL GENOMES

According to currently available information thirty one bifidobacterial strains have had their genomes sequenced to completion (Table 1.1). The first bifidobacterial genome sequenced was that of *B. longum* (subsp. *longum*) NCC2705 (Schell *et al.*, 2002), which was shown to contain a genome with a size of 2.26 Mb. All completely sequenced genomes exhibit a GC content of 58-61 %, and range in size from 1.9 to over 2.8 Mb, with *B. longum* subsp. *infantis* ATCC 15697 possessing the largest genome (2.83 Mb) (Table 1.1). In addition to the thirty one completed bifidobacterial genome sequences, an additional twelve bifidobacterial draft genome sequences are currently listed in the database of the National Centre of Biotechnology Information (NCBI; dated December 1, 2013).

Table 1.1 Completely sequenced bifibifidobacterial genomes

Species	Genome size (Mb)	GC content %	Genes	Proteins	Source	Genbank No.	Reference
Bifidobacterium adolescentis ATCC 15703	2.09	59.2	1,702	1,632	Human GIT	NC_008618	Unpublished
Bifidobacterium animalis subsp. *lactis* V9	1.94	60.5	1,636	1,572	Infant GIT	NC_017217	(Sun *et al.*, 2010)
Bifidobacterium animalis subsp. *animalis* ATCC 25527	1.93	60.5	1,601	1,538	Human GIT	NC_017834	Unpublished
Bifidobacterium animalis subsp. *lactis* AD011	1.93	60.5	1,603	1,527	Infant faeces	NC_011835	(Kim *et al.*, 2009)
Bifidobacterium animalis subsp. *lactis* B420	1.94	60.5	1,625	1,561	Human GIT	NC_017866	(Stahl & Barrangou, 2012)
Bifidobacterium animalis subsp. *lactis* BB-12	1.94	60.5	1,706	1,642	Human GIT	NC_017214	(Garrigues *et al.*, 2010)
Bifidobacterium animalis subsp. *lactis* ATCC 27673	1.96	60.6	1,620	1,556	Human GIT	NC_022523	(Loquasto *et al.*, 2013)
Bifidobacterium animalis subsp. *lactis* BLC1	1.94	60.5	1,607	1,518	Human GIT	NC_017216	(Bottacini *et al.*, 2011)
Bifidobacterium animalis subsp. *lactis* Bi-07	1.94	60.5	1,661	1,597	Human faeces	NC_017867	(Stahl & Barrangou, 2012)
Bifidobacterium animalis subsp. *lactis* Bl-04	1.94	60.5	1,631	1,567	Human faeces	NC_012814	(Barrangou *et al.*, 2009)
Bifidobacterium animalis subsp. *lactis* Bl12	1.94	60.5	1,607	1,518	Human colon	NC_021593	(Milani *et al.*, 2013)
Bifidobacterium animalis subsp. *lactis* CNCM I-2494	1.94	60.5	1,724	1,660	Human faeces	NC_017215	(Chervaux *et al.*, 2011)
Bifidobacterium animalis subsp. *lactis* DSM 10140	1.94	60.5	1,628	1,565	Infant GIT	NC_012815	(Barrangou *et al.*, 2009)
Bifidobacterium asteroides PRL2011	2.17	60.1	1,731	1,658	Honeybee	NC_018720	(Bottacini *et al.*, 2012)
Bifidobacterium bifidum BGN4	2.22	62.6	1,902	1,834	Human faeces	NC_017999	Unpublished
Bifidobacterium bifidum PRL2010	2.21	62.7	1,791	1,706	Infant faeces	NC_014638	(Turroni *et al.*, 2010)
Bifidobacterium bifidum S17	2.19	62.8	1,845	1,783	Infant faeces	NC_014616	Unpublished
Bifidobacterium breve ACS-071-V-Sch8b	2.33	58.7	2,011	1,826	Human GIT	NC_017218	Unpublished
Bifidobacterium breve UCC2003	2.42	58.7	1,985	1,854	Infant faeces	NC_020517	(Maze *et al.*, 2007)
Bifidobacterium dentium Bd1	2.64	58.5	2,195	2,127	Dental caries	NC_013714	(Ventura *et al.*, 2009)
Bifidobacterium longum subsp. *longum* JCM 1217	2.39	60.3	2,009	1,924	Human faeces	NC_015067	(Fukuda *et al.*, 2011)
Bifidobacterium longum subsp. *longum* JDM301	2.48	59.8	2,035	1,958	Human faeces	NC_014169	(Wei *et al.*, 2010)
Bifidobacterium longum subsp. *longum* KACC 91563	2.4	59.8	2,350	1,985	Infant faeces	NC_017221	(Ham *et al.*, 2011)
Bifidobacterium longum NCC2705	2.26	60.1	1,799	1,728	Human GIT	NC_004307	(Schell *et al.*, 2002)
Bifidobacterium longum DJO10A	2.39	60.1	2,073	2,001	Human GIT	NC_010816	(Lee *et al.*, 2008)

Species	Genome size (Mb)	GC content %	Genes	Proteins	Source	Genbank No.	Reference
Bifidobacterium longum subsp. *infantis* ATCC 15697	2.83	59.9	2,588	2,416	Infant GIT	NC_011593	(Sela *et al.*, 2008)
Bifidobacterium longum subsp. *infantis* 157F	2.41	60.1	2,070	1,999	Infant GIT	NC_015052	(Fukuda *et al.*, 2011)
Bifidobacterium longum subsp. *longum* BBMN68	2.27	59.9	1,876	1,804	Human GIT	NC_014656	(Hao *et al.*, 2011)
Bifidobacterium longum subsp. *longum* F8	2.38	59.9	1,743	1,681	Human GIT	NC_021008	Unpublished
Bifidobacterium thermophilum RBL67	2.29	60.1	1,904	1,845	Infant faeces	NC_020546	Unpublished

1.2.2 BIFIDOBACTERIAL PLASMIDS

Approximately 20 % of tested bifidobacterial strains contain detectable plasmids, which are typically small to medium-sized, ranging from 1.5 kb to 15 kb (Cronin *et al.*, 2011; Park *et al.*, 1997a; Sgorbati *et al.*, 1982). Bifidobacterial megaplasmids, i.e. extra-chromosomal genetic elements of 100 kb or more in size, have also been reported from porcine isolates for certain *B. pseudolongum* subsp. *globosum* although they have not been characterised further (Simpson *et al.*, 2003). For a compilation of completely sequenced bifidobacterial plasmids see Table 1.2 (NCBI; dated December 1, 2013).

Table 1.2 Sequenced bifibifidobacterial plasmids

Species	Plasmid	Accession number	Reference
B. pseudocatenulatum	p4M	NC_003527.1	(Gibbs *et al.*, 2006)
B. catenulatum	pBC1	NC_007068.1	(Alvarez-Martin *et al.*, 2007)
B. pseudolongum subp. *globosum*	pASV479	NC_010877.1	(Sangrador-Vegas *et al.*, 2007)
B. bifidum	pBIF10	DQ093580	Unpublished
	PB80	NC_011332.1	(Shkoporov *et al.*, 2008)
B. asteroides	pCIBAO89	NC_010908.1	(Cronin *et al.*, 2007)
	pAP1	Y11549	Unpublished
B. breve	pCIBB1	NC_002133.1	(O'Riordan & Fitzgerald, 1999)
	pNBB1	E17316	(Bourget *et al.*, 1993)
	pB21A	NC_010930.1	(Shkoporov *et al.*, 2008)
B. kashiwanohense	pBBKW-1	AB713428	(Takahata *et al.*, 2013)
	pBBKW-2	NC_021876.1	(Takahata *et al.*, 2013)
B. longum	BLNIAS-p2	NC_017222.1	(Ham *et al.*, 2011)
	pDOJH10L	NC_004252.1	(Lee & O'Sullivan, 2006)
	pKJ50	NC_004978.1	(Park *et al.*, 1999)
	pNAC2	NC_004769.1	(Corneau *et al.*, 2004)
	pB44	NC_004443.1	(Shkoporov *et al.*, 2008)
	pTB6	NC_006843.1	(Tanaka *et al.*, 2005)
	pKJ36	NC_002635.1	(Park *et al.*, 1997a)
	pMG1	NC_006997.1	Unpublished
	p6043B	NC_010861.1	Unpublished
	pBLO1	NC_004943.1	(Schell *et al.*, 2002)
	pNAC1	NC_004779	(Corneau *et al.*, 2004)
	pNAL8M	AM183144.1	(Guglielmetti *et al.*, 2007)
	pBIFA24	NC_010164.1	(Park *et al.*, 2008)
	p6043A	DQ458910	Unpublished
	pNAL8L	AM183145.1	(Guglielmetti *et al.*, 2007)
	pNAC3	NC_004768	(Corneau *et al.*, 2004)
	pDOJH10S	NC_004253.1	(Lee & O'Sullivan, 2006)
	pSPO2	GU256055.1	Unpublished
	pMB1	X84655	(Rossi *et al.*, 1996)
	pFI2576	NC_011139.1	(Moon *et al.*, 2009)
	BLNIAS-p1	NC_017220	(Ham *et al.*, 2011)
	p157F-NC1	NC_015053	(Fukuda *et al.*, 2011)
	p157F-NC2	NC_015066	(Fukuda *et al.*, 2011)
	pBK283	AB495342.1	(Fukiya *et al.*, 2010)

1.2.3 THETA TYPE REPLICATION

Two main types of plasmid replication exist: rolling circle and theta-type replication. A factor which distinguishes theta-type replication from the rolling circle mechanism is that the former mechanism does not produce single stranded DNA replication intermediates. This literature review will focus mainly on theta-type replication, which distinguishes three subtypes based on mechanistic differences of the replication initiation process. The first of these is dependent on a plasmid-encoded replication initiator, the replication or Rep protein, which binds to a series of direct repeats, also known as iterons, to form an initiation complex at the origin of replication. This complex, in association with DnaA, facilitates the recruitment of a DnaB-DnaC protein complex to this plasmid replication origin, resulting in the strand opening of an AT-rich DNA region, directly adjacent to the initiation complex, thereby initiating leading strand synthesis primed by a DnaG-mediated RNA transcript. The organisation and functioning of this replication subtype resembles that of the origin of chromosomal replication (*OriC)* in *E. coli* (Bramhill & Kornberg, 1988).

The second recognised subtype of theta-type replication initiation is not dependent on a plasmid-encoded initiator protein, yet is instead dependent on host DNA polymerase I. In this mechanism the initial strand opening event occurs by a strand displacement reaction at the origin due to synthesis of an RNA transcript. This RNA transcript is processed by Rnase H, and this processed RNA is then used for leading strand synthesis which is initially performed by DNA polymerase I, but which is replaced by DNA polymerase III holoenzyme for DNA elongation purposes. A prime example of this mechanism of replication is that observed for the ColE1 plasmid (Kingsbury & Helinski, 1970).

The third subtype of replication initiation found in theta-type replicons requires host DNA polymerase I and a Rep protein, and is the mechanism of choice for ColE2 and ColE3 type replicons (Yasueda *et al.*, 1989). In this case the Rep protein binds to the replication origin, which is located downstream of the Rep-encoding gene. This interaction results in the production of an RNA primer, catalysed by the host RNA polymerase, and is responsible for initiation of leading strand synthesis by DNA polymerase I (Takechi *et al.*, 1995).

Following initiation by either of the previously mentioned three mechanisms, theta-type replication can occur either bidirectionally (Prescott & Kuempel, 1972) or unidirectionally (Takechi & Itoh, 1995), with most theta-type plasmids replicating unidirectionally. Termination of replication occurs when the replication fork composed of the leading strand and lagging strand comes in contact with a termination sequence. The termination sites specifically interact with a replication terminator protein and this complex halts leading strand synthesis (Mulugu *et al.*, 2001).

1.2.4 E. COLI/BIFIDOBACTERIUM SHUTTLE VECTORS

A number of *E. coli/Bifidobacterium* shuttle vectors have been constructed to date (Table 1.3), and such shuttle vectors are crucially important to facilitate genetic manipulation of bifidobacteria. Examples are the pBC1.2 shuttle vector, constructed based on the *B. catenulatum* pBC1-based replicon, and functional in several *Bifidobacterium* species, including *B. breve* UCC2003 and *B. pseudocatenulatum* M115 (Alvarez-Martin *et al.*, 2007), and the pPKCM1 shuttle vector, which is based on a plasmid, pCIBAO89, from *B. asteroides*, and which replicates in *B. animalis* subsp. *lactis*, *B. longum* NCIMB 8809, *B. pseudolongum* NCIMB 2244, *B.*

pseudolongum subsp. globosum JCM 5820, *B. pseudocatenulatum* LMG 10505 and *B. dentium* NCFB 2843 (Cronin *et al.*, 2007).

Another example is the replicon of the cryptic *B. longum* plasmid pBLES100, which when cloned into the *E. coli* vector pBR322 generated a shuttle vector pBLES100 that was capable of transforming *B. longum* 105-A and 108-A (Matsumura *et al.*, 1997). Various other examples of successful *E. coli/Bifidobacterium* shuttle vectors have been published (Klijn *et al.*, 2006; Missich *et al.*, 1994; Rossi *et al.*, 1996; Rossi *et al.*, 1998; Sangrador-Vegas *et al.*, 2007; Shkoporov *et al.*, 2008), and for more information on such *E. coli/Bifidobacterium* shuttle vectors see Table 1.3.

Many *E. coli/Bifidobacterium* shuttle vectors constructed to date have the ability to replicate in several bidfidobacterial strains; however, experiments to test the replication of some of these plasmids in strains of the Gram positive genera, *Lactobacillus* and *Lactococcus*, have not produced transformants. For example the shuttle vector pDOJHR, which is based on a plasmid found in *B. longum* DOJH10S, was shown to be only capable of replicating in *B. longum* VMK44, thus exhibiting a very limited host range (Lee & O'Sullivan, 2006;) . Another plasmid with a relatively limited host range is shuttle vector pPKCM1 which appears to be unable to transform *L. lactis* and it was thus concluded that this plasmid is restricted to certain Gram positive bacteria (Cronin *et al.*, 2007) suggesting that bifidobacterial plasmid replication functions may be limited to this genus or perhaps high GC Gram positive bacteria.

Table 1.3 *E. coli/Bifidobacterium* shuttle vectors constructed to date

Strain origin of plasmid	Cryptic bif plasmid	*E.coli* vector	Name shuttle vector	Transformable strains	Reference
B. catenulatum L48	pBC1	pBIF	pBC1.2	*B. breve* UCC2003	(Alvarez-Martin *et al.*, 2007)
				B. pseudocatenulatum M115	
B. asteroides	pCIBAO89	pBlueCm	pPKCm1	*B. animalis* subsp. *lactis*,	(Cronin *et al.*, 2007)
				B. longum NCIMB 8809	
				B. longum NCIMB 8809	
				B. pseudolongum NCIMB 2244,	
				B. globosum JCM5820,	
				B. pseudocatenulatum LMG10505	
				B. dentium NCFB 2843	
B. longum NCC293	pNCC293	pMDY23	pDP870	*B. longum* NCC2705	(Klijn *et al.*, 2006)
B. longum DJO10A	pDOJH10S	pACYC184	pDOJHR	*B. longum* VMK44	(Lee & O'Sullivan, 2006)
B. longum 105-A	pBLES100	pBR322	pBLES100	*B. longum* 105-A	(Matsumura *et al.*, 1997)
				B. longum 108-A	
B. longum 2577	pRM1	pGEM-5Z7	pRM2	*B. longum* 2577	(Missich *et al.*, 1994)
B. longum B2577	pMB1	pBluescriptll KS	pDG7	*B. animalis* MB209	(Rossi *et al.*, 1996)
	pMB1	pBluescriptII KS	pLF5	*B. bifidum* MB254	(Rossi *et al.*, 1998)
				B. animalis MB209	
				B. infantis MB263	
				B. longum MB219	
				B. magnum MB267	
		pBluescriptII KS	pCLJ15	*B. bifidum* MB254	(Rossi *et al.*, 1998)
				B. animalis MB209	
				B. infantis MB263	

Strain origin of plasmid	Cryptic bif plasmid	*E. coli* vector	Name shuttle vector	Transformable strains	Reference
				B. longum MB219	
				B. magnum MB267	
		pBluescriptII KS	pSPEC1	*B. bifidum* MB254	(Rossi *et al*., 1998)
				B. animalis MB209	
				B. infantis MB263	
				B. longum MB219	
				B. magnum MB267	
B. pseudolongum subsp.*globosum* DPC479	pASV479	pBluescriptII KS	pASV480	*B. breve* NCIMB 8807	(Sangrador-Vegas *et al*., 2007)
B. longum	pB44	pE194	pSUW1/114	*B. bifidum* ATCC15696	(Shkoporov *et al*., 2008)
			pSUW53/114	*B. bifidum* ATCC15696	(Shkoporov *et al*., 2008)
		pUC19E	pSUW64/123	*B. bifidum* ATCC15696	(Shkoporov *et al*., 2008)
				B. breve UCC2003	
B. bifidum	pB80	pBluescript II KS	pEESH80	*B. breve* UCC2003	(Shkoporov *et al*., 2008)
			pCESH80	*B. breve* UCC2003	(Shkoporov *et al*., 2008)

1.3 PROBIOTICS, PREBIOTICS AND SYNBIOTICS

Probiotics are defined as live microbial food ingredients that have a beneficial impact on human health (Salminen *et al.,* 1998). The term probiotic is based on two Greek words, meaning 'for life', and the concept of probiotics evolved at the turn of the 20th century due to a hypothesis put forward by the Russian scientist Elie Metchnikoff (Bibel, 1988). He believed that Bulgarian peasants lived long and healthy lives due to the fact that they consumed fermented milk products.

Bifidobacteria are commercially exploited as health-promoting or probiotic ingredients in certain functional foods due to their perceived role in preventing disease and enhancing gastrointestinal health. The reported beneficial effects that specific strains of bifidobacteria may elicit on the host include, among many others, (i) inhibition of pathogens through acetate production which is thought to improve intestinal defence by preventing the passage of pathogenic bacteria from the gut lumen to the blood (Fukuda *et al.,* 2011), (ii) alleviation of lactose intolerance by consumption of a combination of *Lactobacillus casei* Shirota and *B. breve* Yakult (Almeida *et al.,* 2012), and (iii) reduction of serum cholesterol levels by the administration of a cocktail of bifidobacteria (Bordoni *et al.,* 2013).

Various microbes other than bifidobacteria have been associated with probiotic activities. For example, it was recently shown that *Lactobacillus reuteri* DSM 17938, in conjunction with rehydration therapy, is efficient in the treatment of acute diarrhoea (Francavilla *et al.,* 2012). *Propionibacterium freudenreichii* has also been indicated as a probiotic strain because of its immunomodulatory and growth-promoting effects in pigs (Cousin *et al.,* 2012).

The search for new (bifidobacterial) probiotics is on-going (for reviews see (Collado *et al.*, 2012; Jankovic *et al.*, 2010; Quigley, 2010)). For example, in a recent study four bifidobacterial strains (*B. longum* 51A, *B. breve* 1101A, *B. pseudolongum* 1191A and *B. bifidum* 1622A), isolated from faeces of healthy children, were compared based on their growth rates, aerotolerance, antagonistic activity against pathogens, antimicrobial susceptibility profile and cell wall hydrophobicity (Souza *et al.*, 2013). Based on these analyses *B. longum* 51A was shown to possess the best potential for probiotic use as it did not appear to contain any antibiotic resistance determinants, while it exhibited the highest growth rate and the best ability to produce antagonistic substances against a range of pathogenic microorganisms (Souza *et al.*, 2013).

The term prebiotic was first coined nearly twenty years ago (Gibson & Roberfroid, 1995). These authors defined prebiotics as "non-digestible food ingredients that beneficially affect the host by selectively stimulating the growth and/or activity of one or a limited number of bacteria in the colon". Synbiotics are defined as "a mixture of probiotics and prebiotics that beneficially affects the host by improving the survival and implantation of live microbial dietary supplements in the gastrointestinal tract, by selectively stimulating the growth and/or by activating the metabolism of one or more limited number of health-promoting bacteria and thus improving host welfare" (Gibson & Roberfroid, 1995).

It has been demonstrated that the proliferation of gut-derived probiotic bacteria such as bifidobacteria is selectively stimulated by various dietary carbohydrates, which for this reason are considered to be prebiotics (Macfarlane *et al.*, 2008). Various human clinical trials have been performed to analyse the efficacy of prebiotics. For example, regular ingestion of lactulose (10 g per day) was shown to result in an

increase in bifidobacterial numbers present in the faecal microbiota of participants, while having little or no effect on total bacterial populations (Tuohy *et al.*, 2002). Similarly, ingestion of 'Vivinal' galacto-oligosaccharides (GOS; 2 g twice daily) was demonstrated to significantly increase population levels of bifidobacteria in faecal samples of subjects over 50 (Walton *et al.*, 2012), while polydextrose was shown to result in a decrease in numbers of *Bacteroides* species, and a numerical increase in representatives of *Lactobacillus* and *Bifidobacterium* species (Jie *et al.*, 2000). Furthermore, lactitol (ingested twice a day, total daily dose of 10 g) in conjunction with the administration of the probiotic *Lactobacillus acidophilus* NCFM (daily dose of 2×10^{10} cells) was also shown to result in an increase in bifidobacterial levels (Bjorklund *et al.*, 2012).

1.4 CARBOHYDRATE METABOLISM IN BIFIDOBACTERIA

The normal human diet contains a non-digestible (i.e. resistant to human hydrolytic enzymes) carbohydrate fraction, and these non-digestible dietary components include plant cell wall polysaccharides, such as arabinan and a particular starch fraction, also known as resistant starch, (as it reaches the large intestine without being broken down) (Grabitske & Slavin, 2009; Lattimer & Haub, 2010; Topping & Clifton, 2001). It has been estimated that approximately 20-60 g of dietary carbohydrates reach the colon daily, having escaped or being insensitive to digestion by host enzymes (Cummings & Macfarlane, 1991; Silvester *et al.*, 1995). Non-digestible carbohydrates present in the diet not only have a profound effect on the composition of the human GIT microbiota, but also influence the total bacterial load in the gut and the concentration of associated metabolic products (Stephen *et al.*, 1987).

As mentioned above, bifidobacteria operate a unique hexose metabolism, the so-called bifid-shunt, a metabolic pathway that employs the signature enzyme fructose-6-phosphate-phosphoketolase (F6PPK), converting a range of monosaccharides to mainly acetic and lactic acid ((Scardovi & Trovatelli, 1965) see Fig. 1.1). This pathway provides bifidobacteria an advantage over other bacteria in that it allows the production of more energy in the form of ATP from carbohydrates than alternative fermentative pathways present in other bacteria such as lactic acid bacteria. The bifidobacterial pathway yields 2.5 ATP molecules from 1 molecule of fermented glucose, as well as 1.5 molecules of acetate and 1 molecule of lactate (Palframan *et al.*, 2003), although this lactate to acetate ratio may vary depending on the carbon source and growth phase, while it is also species dependent (Palframan *et al.*, 2003).

Approximately 8 % of the average bifidobacterial genome is dedicated to carbohydrate metabolism, and half of these genes are putatively responsible for carbohydrate uptake, which occurs mainly by means of ATP-binding cassette transporters (ABC-type transporters). Just twenty five genes present on the genome of *B. bifidum* PRL2010 were predicted to encode components of transport systems dedicated to carbohydrate uptake (Turroni *et al.*, 2012c), and it was observed that in comparison to other bifidobacteria (with the exception of *B. animalis* subsp. *lactis* Bb12) this strain appears to be able to metabolise only a small number of carbohydrates. Of these twenty five genes, eight represented members of the ABC-type transport family for carbohydrates such as fructose, glucose, ribose, galactose and turanose.

In addition to ABC-type carbohydrate transporters, bifidobacteria also employ proton motive force-driven permeases, proton symporters, and phosphoenolpyruvate-dependent phosphotransferase systems (PEP-PTSs) for

carbohydrate internalisation (Maze *et al.*, 2007; Pokusaeva *et al.*, 2011a; Schell *et al.*, 2002). The PEP-PTS system is only found in bacteria (although not in all bacteria), being responsible for the transport and phosphorylation of various carbohydrates, while also playing a crucial role in the global regulation of carbohydrate metabolism and various other physiological processes (Postma *et al.*, 1993). PEP is used both as an energy source as well as a phosphoryl donor, where a phosphoryl group of PEP is transferred via four proteins to the membrane components of the PTS, forming the PTS phosphorylation cascade (Deutscher *et al.*, 2006). A more comprehensive discussion of PEP-PTS systems and their role in global regulation, and their presence in bifidobacteria will follow in the next section.

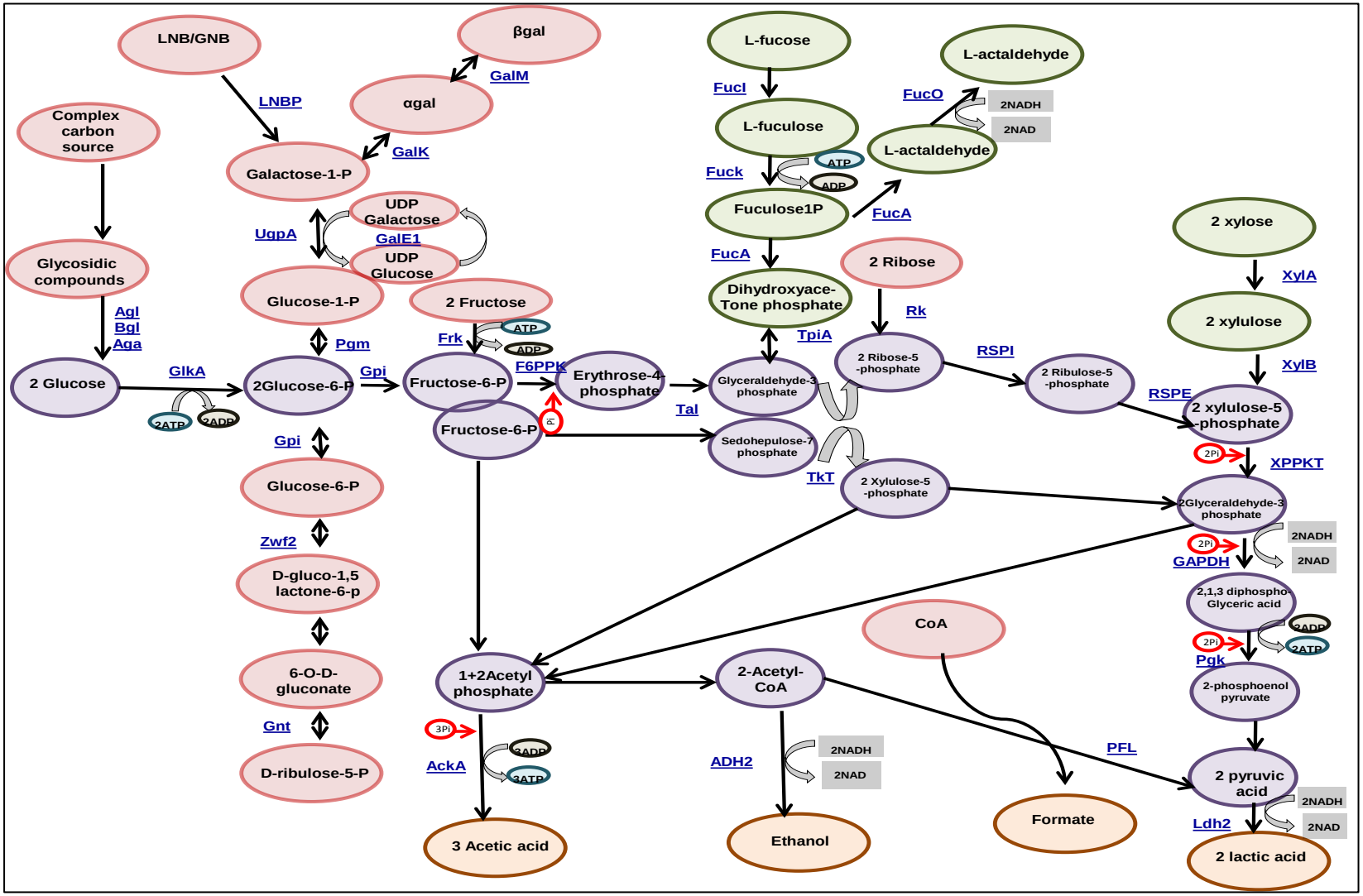
LNB/GNB
βgal
GalM
αgal
LNBP
Complex carbon source
GalK
Galactose-1-P
UDP Galactose
UgpA
GalE1
UDP Glucose
Glycosidic compounds
Glucose-1-P
2 Fructose
Agl
Bgl
Aga
Pgm
Frk
ATP
ADP
2 Glucose
GlkA
2ATP
2ADP
2Glucose-6-P
Gpi
Fructose-6-P
F6PPK
Erythrose-4-phosphate
Fructose-6-P
Gpi
Glucose-6-P
Zwf2
D-gluco-1,5 lactone-6-p
6-O-D-gluconate
Gnt
D-ribulose-5-P
L-fucose
FucI
L-fuculose
Fuck
ATP
ADP
Fuculose1P
FucA
FucA
Dihydroxyace-Tone phosphate
TpiA
Glyceraldehyde-3 phosphate
Tal
Sedohepulose-7 phosphate
L-actaldehyde
FucO
2NADH
2NAD
L-actaldehyde
2 Ribose
Rk
2 Ribose-5 -phosphate
RSPI
2 Ribulose-5 -phosphate
TkT
2 Xylulose-5 -phosphate
2 xylose
XylA
2 xylulose
XylB
RSPE
2 xylulose-5 -phosphate
2Pi
XPPKT
2Glyceraldehyde-3 phosphate
2Pi
GAPDH
2NADH
2NAD
2,1,3 diphospho-Glyceric acid
2ADP
2ATP
2Pi
Pgk
2-phosphoenol pyruvate
1+2Acetyl phosphate
3Pi
AckA
3ADP
3ATP
3 Acetic acid
2-Acetyl-CoA
ADH2
2NADH
2NAD
Ethanol
CoA
PFL
Formate
2 pyruvic acid
2NADH
2NAD
Ldh2
2 lactic acid

Figure 1.1 Schematic representation of carbohydrate degradation through the “bifid shunt” in bifidobacteria. Abbreviations: AckA, acetate kinase; Adh2, aldehyde-alcohol dehydrogenase 2; Aga, α-galactosidase; Agl, α-glucosidase; Bgl, β-glucosidase; GalE1, UDP-glucose 4-epimerase; GalK, galactokinase; GalM, galactose mutarotase; GAPDH, glyceraldehyde-3-phosphate dehydrogenase C; GlkA, glucokinase; Gnt, 6-phosphogluconate dehydrogenase; Gpi, glucose 6-phosphate isomerase; Frk, fruktokinase; F6PPK, fructose-6-phosphoketolase; FucI, L-fucose isomerase; FucK, L-fuculose kinase; FucA, L-fuculose-1P aldolase; FucO, lactaldehyde reductase; Ldh2, lactate dehydrogenase; LNBP, lacto-N-biose phosphorylase; Pgk, phosphoglyceric kinase; Pgm, phosphoglucomutase; Pfl, formate acetyltransferase; Rk, ribokinase; R5PI, ribose-5-phosphate isomerase; R5PE, ribulose-5-phosphate epimerase; Tal, transaldolase; Tkt, transketolase; TpiA, triosephosphate isomerase; UgpA, UTP-glucose-1-phosphate uridylyltransferase; XPPKT, xylulose-5-phosphate/fructose-6-phosphate phosphoketolase; XylA, xylose isomerase; XylB, xylulose kinase; Zwf2, glucose-6-phosphate 1-dehydrogenase; Pi, phosphate (based on a figure from a previous review (Pokusaeva *et al.*, 2011a)).

1.4.1 PEP-PTS SYSTEM AND CATABOLITE REPRESSION

In many bacteria the carbohydrate flux going through a PEP-PTS system regulates the activity of other non-PTS systems. This regulation is based on the availability of a preferred carbon source in the growth medium, which indirectly inhibits the expression or activity of proteins involved in the utilisation or uptake of other available carbon sources, a global regulatory process commonly referred to as carbon catabolite repression (CCR) (Postma *et al.*, 1993; Saier & Ramseier, 1996). In a recent study by Sela *et al.*, (2012), their findings supported the idea of a carbon catabolite repression system being in operation in *B. longum* subsp. *infantis* ATCC 15697. A mixture of six low molecular weight human milk oligosaccharides (HMOs) induced the transcription of Blon_2335, which encodes a GH95 family α-fucosidase, while it repressed transcription of Blon_0248 and Blon_0426, which represent two GH29 family α-L-fucosidases. It was hypothesised that this repression was due to lacto-N-tetraose as, when cells were grown on this substrate, transcription of Blon_0248 and Blon_0426 was also repressed. Similarly, a lactose-over-glucose preference was observed in *B. longum* NCC2705 (Parche *et al.*, 2006). In this case transcription of *glcP*, which encodes a glucose transporter, is repressed in the presence of lactose.

Further examples of catabolite repression have been described for *B. breve* UCC2003. In the case of ribose metabolism by this bifidobacterial strain it was observed that transcription of the ribose cluster is decreased by two fold when the culture in grown in a mixture of glucose and ribose (Pokusaeva *et al.*, 2010). Furthermore, transcription of the *fos* operon of *B. breve* UCC2003, involved in oligofructose metabolism, was observed to be induced in the presence of

oligofructose or sucrose, but shown to be repressed in the presence of a mixture of glucose and sucrose, or fructose and sucrose (Ryan *et al.*, 2005).

In terms of bifidobacterial PEP-PTS systems, four complete PEP-PTS systems were predicted to be encoded by the *B. breve* UCC2003 genome (Maze *et al.*, 2007), more than most other analysed bifidobacterial genomes. For example *B. longum* subsp. *longum* DJO10A (Lorca *et al.*, 2007) and *B. longum* subsp. *longum* NCC2705 (Parche *et al.*, 2007) both encode just one PEP-PTS system, while *B. animalis* subsp. *lactis* (Barrangou *et al.*, 2009) does not appear to encode a single PEP-PTS system. The finding that some bifidobacteria do not possess a PEP-PTS system, combined with the notion that specific regulatory proteins required for CCR in other bacteria appear to be absent in bifidobacteria (unpublished data), indicated that bifidobacteria must control CCR via an as yet unknown regulatory mechanism. As has been shown in various reports, transcription of the genes encoding proteins involved in the metabolism of a specific carbohydrate is regulated. For example, a recently published transcriptome analysis of *B. animalis* subsp. *lactis* Bl-04 reported on carbohydrate-dependent induction by eleven oligosaccharides and the consequent identification of genetic loci involved in the uptake and metabolism of these saccharidic compounds (Andersen *et al.*, 2013). Other examples will be discussed below, together with the identified regulatory mechanisms.

1.4.2 LACI AND ROK MEDIATED REPRESSION SYSTEMS

LacI-mediated repression was first characterised for the *lac* operon in *E. coli* (Jacob & Monod, 1961). The *lac* operon consists of three genes *lacZ*, *lacY* and *lacA*, which encode a β-galactosidase, a lactose permease and a transacetylase, respectively. In the absence of lactose, the repressor protein also known as LacI binds to target

sequences called operator sequences, which prevents the RNA polymerase from transcribing the genes present in the operon. When lactose is present as the main source of energy, some of the lactose is converted into allolactose, which, by its physical association with LacI, changes the conformation of this repressor and prevents LacI binding to its cognate operator sequences, thereby allowing transcription of the *lac* operon. LacI/GalR family regulators share similar structures and regulate transcription in a similar manner (Wintjens & Rooman, 1996). LacI/GalR family regulators are composed of both DNA-binding and regulatory domains, the latter possessing two pivotal roles: (i) to receive and transmit the signal via effector binding, and (ii) to mediate homodimer formation/disassembly.

Various members of the LacI/GalR family have been identified in bifidobacteria and in several cases their role in carbohydrate-dependent regulation of gene expression has been demonstrated. An example of LacI-type mediated transcriptional control in bifidobacteria is the *galACDEGR* locus of *B. breve* UCC2003, which is involved in utilisation of galactan, a plant-derived polysaccharide. The *galA* gene encodes a galactanase which degrades the galactan polymer predominantly into galactotriose/biose, *galC*, *galD* and *galE*, which specifies a presumed ABC-type uptake system for the GalA-produced galacto-oligosaccharides, *galG* encodes a β-galactosidase, while *galR* encodes a LacI-type DNA-binding protein. Transcription of the *galCDEG* operon and the monocistronic *galA* is controlled by GalR, and it was observed that this LacI-type repressor responds to the carbohydrate effector galactobiose: in the absence of galactobiose, GalR binds to its operator sites thereby blocking transcription of the *gal* genes, while in the presence of this effector GalR is unable to bind to these DNA sequences (O'Connell Motherway *et al.*, 2011a).

Other such carbohydrate metabolic systems regulated by LacI-type transcriptional repressors would be that of cellodextrin and ribose metabolism in *B. breve* UCC2003. Cellodextrin utilisation by *B. breve* UCC2003 requires the products of the *cldEFGC* gene cluster (Pokusaeva *et al.*, 2011b). The *cldE, cldF* and *cldG* genes encode a putative cellodextrin-binding protein and two cellodextrin permease proteins, respectively, while *cldC* specifies a β-glucosidase. Through electrophoretic mobility shift assay (EMSA) analysis it was shown that this cluster is negatively regulated by a LacI-type regulator encoded by *cldR*. Ribose metabolism by *B. breve* UCC2003 involves the gene products of the *rbsACBDK* gene cluster. The *rbsA*, *rbsC* and *rbsB* genes, encode the ribose uptake system. A putative ribose mutarotase is encoded by *rbsD* and a putative ribokinase is encoded by the *rbsK* gene. Similar to the *cld* cluster, the transcription of the *rbs* cluster was shown to be negatively regulated by a dedicated LacI-type regulator encoded by *rbsR* (Pokusaeva *et al.*, 2010).

Repressor open reading frame kinase (ROK) regulators were originally described approximately twenty years ago and include transcriptional repressors, sugar kinases and as yet uncharacterised open reading frames (Titgemeyer *et al.*, 1994). ROK repressors and kinases can be distinguished based on their N-terminal sequence: ROK kinases possess a conserved N-terminal ATP-binding motif, while ROK repressors contain an additional N-terminal DNA-binding domain and a C-terminal sugar-binding domain (Titgemeyer *et al.*, 1994). The ROK protein family is characterised by the presence of the PF00480 domain that belongs to the Actin-ATPase clan in the Pfam database (Punta *et al.*, 2012). Mlc from *E. coli* is a ROK transcriptional repressor which controls the expression of genes encoding enzymes of the phosphotransferase system (PTS), the Mlc ROK repressor from *E. coli* also

controls the transcription of *malT*, which is the global activator of the *mal* regulon (Schiefner *et al.,* 2005). The latter study describes the first structure of a ROK family protein. Examples of regulators in this family that have been characterised are NagC from *E. coli* (Plumbridge, 2001; Plumbridge *et al*., 1993), XylR from *Bacillus subtilis* (Dahl *et al.*, 1994; Scheler & Hillen, 1994) and CsnR from *Streptomyces lividans* (Dubeau *et al*., 2011).

1.4.3 GLYCOSYL HYDROLASES (GHs) IN BIFIDOBACTERIA

GHs (EC 3.2.1.-) are a widespread group of enzymes which hydrolyse the glycosidic bond between two or more monosaccharides, or between a carbohydrate and a non-carbohydrate moiety (Cantarel *et al.*, 2009). The first three digits of the assigned EC number indicate enzymes that hydrolyse O-glycosyl linkages, whereas the last number indicates the substrate and sometimes reflects the molecular mechanism. Classification of GHs into families is based on amino acid sequence similarities and was devised in 1991 (Henrissat, 1991). GHs which exhibit significant sequence similarities to each other were assigned to a particular GH family. Typically, hydrolysis of the glycosidic bond is catalysed by two amino acid residues of the enzyme to provide a general acid (proton donor) and a nucleophile/base (Davies & Henrissat, 1995). Classification of glycoside hydrolases is available at http://www.cazy.org/Glycoside-Hydrolases.html (Cantarel *et al.*, 2009).

Humans do not possess enzymes required for the degradation of prebiotic carbohydrates, e.g. fructo-oligosaccharides (FOS), galacto-oligosaccharides (GOS), xylo-oligosaccharides, inulin, arabinoxylan and lactulose (Guarner & Malagelada, 2003). Gut commensals such as bifidobacteria utilise GHs to degrade dietary carbohydrates, although other enzyme activities, such as polysaccharide lyases and

carbohydrate esterases, may also be utilised (Flint *et al.*, 2012; Pokusaeva *et al.*, 2011a). For a full list of putative glycosyl hydrolases present in currently sequenced bifidobacteria, see Tables 1.4-1.4.4 (CAZY; dated December 1, 2013). It is not possible to cover all characterised or predicted carbohydrate-degrading enzymes from bifidobacteria in any significant detail, and therefore in the following sections we will focus on a number of representative GH that are mostly related to the metabolism of plant-derived carbohydrates.

Table 1.4 Bifdobacterial glycosyl hydrolases

		Cazy family	*B. longum* NCC2705	*B. longum* DJO10A	*B. longum* subsp. *infantis* ATCC 15697	*B.longum* subsp. *infantis* 157F	*B. animalis* subsp. *lactis* V9
EC Number	Enzyme name						
3.2.1.21	β-glucosidase/β-D-glucoside glucohydrolase	GH1/3	2	3	3	4	3
3.2.1.23	β-galactosidase	GH1/2/35/42	2	2	5	4	4
3.2.1.10	Oligo-1,6-glucosidase	GH13	1	1		2	1
3.2.1.20	α-glucosidase	GH4/13/31/63/97	3	1		1	1
2.4.1.7	Sucrose phosphorylase	GH13/57/119	1	1	1	1	1
3.2.1.1	α-amylase/neopullalanase	GH13/57		1	4	1	2
2.4.1.18	α-1,4-glucan branching enzyme	GH13/57	1		3	3	1
3.2.1.68	Glycogen operon protein	GH13	2				2
3.2.1.54	Cyclo-maltodextrinase	GH13/57	1				
3.2.1.26	β-fructosidase	GH32	1		2	1	
3.2.1.18	Sialidase	GH33/34/83			2		

3.2.1.22	α-galactosidase	GH4/27/36/57/97	2	2	1	2	3
3.2.1.24	α-mannosidase	GH31/38/47/92	3	1	2	2	
2.4.1.25	4-α-glucanotransferase	GH13/57/77	2	2	2	2	2
3.2.1.96	Endo-β-N-acetyl glucosinaminidase	GH18/73/85/111			1	1	
3.2.1.52	β -N-acetylglucosaminidase	GH20/84/85	1	1	3	1	
3.2.1.140	Lacto-N-biose phosphorylase	GH20/112		1		2	
3.2.1.52	β -N-acetyl hexosaminidase	GH3/20	3			1	1
3.2.1.55	α-L-Arabinosidase/arabinofuranisidase	GH3/43/51	10	4	1	5	2
3.2.1.37	β-xylosidase	GH1/3/30/39/43/52/116	1	5		3	1
3.2.1.41	Pullulanase	GH13		1			2
3.2.1.17	1,4- β -N-acetylmuramidase	GH25/108			4	1	2
3.2.1.25	β-mannosidase	GH1/2/5					1
3.2.1.177	α-xylosidase	GH31	1			2	
3.2.1.8	Endo-1,4-βxylanase	GH5/8/10/11/30/43/51	1		1	3	
3.2.1.17	Putative phage lysin	GH25				1	
3.2.1.53	β-N-acetyl galactosaminidase	GH101				1	
3.2.1.93	Trehalose-6-phosphate hydrolase	GH13		1			1
3.2.1.89	Arabinogalactan-endo-1,4-β-galactosidase	GH53	1	1		1	
3.2.1.51	α-L-fucosidase	GH29			2		
3.2.1.58	Glucan-1,3-β-glucosidase	GH3/5/17	1		1	1	
3.2.1.97	Endo-α-N-acetylgalactosaminidase	GH109			1		
3.2.1.75	Glucan-endo-1,6-β-glucosidase	GH5/30					1
5.4.99.15	Malto oligosyl trehalose synthase	GH13		1			1
NA	Unknown	GH5	1	1	1	1	
NA	Unknown	GH13	1	1		3	
NA	Unknown	GH23	3	3		2	2

NA	Unknown	GH30	1			1	1
NA	Unknown	GH43	2	1		2	
NA	Unknown	GH101	1				
NA	Unknown	GH112	1		1		
NA	Unknown	GH121	1			1	
NA	Unknown	GH125	1	1		1	
NA	Unknown	GH127	2	1		2	1
NA	Unknown	GH129	1	1	1	1	
NA	Unknown	GH120		1			
NA	Unknown	GH32		1			
NA	Unknown	GH95			1		
NA	Unknown	NC*			1		
NA	Unknown	GH65				1	
NA	Unknown	GH94					1
NA	Unknown	GH25	1		2	1	
Total			56	40	46	62	38

NC* indicates Not Characterised

Table 1.4.1 Bifdobacterial glycosyl hydrolases

		Cazy family	*B. animalis* subsp. *lactis* AD011	*B.animalis* subsp. *lactis* B420	*B. animalis* subsp. *lactis* BB-12	*B. breve* ACS-071-V-Sch8b	*B. breve* UCC2003	*B. bifidum* BGN4
EC Number	Enzyme name							
3.2.1.74	Cellodextrinase	GH1/5/9					1	
3.2.1.21	β-glucosidase/β-D-glucoside glucohydrolase	GH1/3	2	3	3	6	3	
3.2.1.23	β-galactosidase	GH1/2/35/42	5	4	4	6	10	5
3.2.1.10	Oligo-1,6-glucosidase	GH13		1	1			
3.2.1.20	α-glucosidase	GH4/13/31/63/97	3	1	1		7	1
2.4.1.7	Sucrose phosphorylase	GH13/57/119	1	2	2	1	1	1
3.2.1.1	α-amylase/neopullalanase	GH13/57	2	2	3	6	3	
2.4.1.18	α-1,4-glucan branching enzyme	GH13/57	1	1	1	1	1	1
3.2.1.68	Glycogen operon protein	GH13					2	
3.2.1.54	Cyclomaltodextrinase	GH13/57					1	1
3.2.1.14	Chitinase	GH18/19/48				1	1	
3.2.1.45	Glucosylceramidase	GH3/30	1	2	2		1	
3.2.1.26	β-fructosidase	GH32	1			2	2	
3.2.1.18	Sialidase	GH33/34/83				1	1	2

3.2.1.22	α-galactosidase	GH4/27/36/57/97	3	3	3	1	2	2
3.2.1.24	α-mannosidase	GH31/38/47/92				2	3	
3.2.1.89	Endogalactanase	GH53					1	
2.4.1.230	Kojibiose phosphorylase	GH65				1		
2.4.1.25	4-α-glucanotransferase	GH13/57/77	2	2	2	2	2	1
3.2.1.96	Endo- β-N-acetyl glucosinaminidase	GH18/73/85/111					1	
3.2.1.52	β-N-acetylglucosaminidase	GH20/84/85				1		
3.2.1.140	Lacto-N-biose phorylase	GH20/112					1	2
3.2.1.52	β -N-acetyl hexosaminidase	GH3/20	1	1	1		2	5
3.2.1.55	α -L-Arabinosidase/arabinofuranosidase	GH3/43/51	3	1	2			1
3.2.1.37	β -xylosidase	GH1/3/30/39/43/52/116	1	3	2			
3.2.1.41	Pullulanase	GH13	1	4	3	1		
3.2.1.17	1,4-β-N-acetylmuramidase	GH25/108	2	2				1
3.2.1.25	β-mannosidase	GH1/2/5						
3.2.1.177	α-xylosidase	GH31						
3.2.1.8	Endo-1,4-β-xylanase	GH5/8/10/11/30/43/51		1	1	2		1
3.2.1.93	Trehalose-6-phosphate hydrolase	GH13			2			
3.2.1.68	Isoamylase	GH13			2			
3.2.1.89	Arabinogalactan-endo-1,4-β-galactosidase	GH53				1		
2.4.1.211	1,3-β-galactosyl-N-acetylhexosamine phosphorylase	GH112				1		
3.2.1.86	6-phospho-β-glucosidase	GH1						1
3.2.1.51	α-L-fucosidase	GH29						2
3.2.1.35	Hyaluronon glucosaminidase	GH84						3
3.2.1.50	α-N-acetylglucosaminidase	GH89						1
3.2.1.58	Glucan-1,3-β-glucosidase	GH3/5/17					1	
3.2.1.97	Endo-α-N-acetylgalactosaminidase	GH109						1

2.4.1.20	Cellibiose phosphorylase	GH94		1				
2.4.1.25/3.2.1.33	Glycogen debranching Enzyme	GH13	2	2		1		2
3.2.1.40	α-L-rhamnosidase	GH78/106	1					
3.2.1.78	Mannon-endo-1,4-β-mannosidase	GH2/5	1		1			
3.2.1.16	Trehalose synthase	GH13		1				
NA	Unknown	GH5				2	1	
NA	Unknown	GH13	1			2		1
NA	Unknown	GH23	2	2	2		2	
NA	Unknown	GH30	1					
NA	Unknown	GH43						1
NA	Unknown	GH125				1	1	
NA	Unknown	GH127	1	1	1			
NA	Unknown	GH129					1	1
NA	Unknown	GH95				1	1	
NA	Unknown	NC				1	1	
NA	Unknown	GH65					1	
NA	Unknown	GH94			1			
NA	Unknown	GH2				1		
NA	Unknown	GH33						1
NA	Unknown	GH36				1		
NA	Unknown	GH25			2	3		3
NA	Unknown	GH53						
Total			38	40	42	49	55	41

NC* indicates Not Characterised

Table 1.4.2 Bifdobacterial glycosyl hydrolases

		Cazy family	*B. bifidum* PRL2010	*B. bifidum* S17	*B. dentium* Bd1	*B. adolescentis* ATCC 15703	*B. thermophilum* RBL67	*B. asteroides* PRL2011
EC Number	Enzyme name							
3.2.1.21	β-glucosidase/β-D-glucoside glucohydrolase	GH1/3	2		13	5	1	3
3.2.1.23	β-galactosidase	GH1/2/35/42	5	4	11	7	1	5
3.2.1.10	Oligo-1,6-glucosidase	GH13			2	1	2	1
3.2.1.20	α-glucosidase	GH4/13/31/63/97	1	1	5	3	2	3
2.4.1.7	Sucrose phosphorylase	GH13/57/119	1	1	2	1	1	2
3.2.1.1	α-amylase/neopullalanase	GH13/57	1	1	1	3	1	1
2.4.1.18	α-1,4-glucan branching enzyme	GH13/57	1	1	1	1	1	
3.2.1.68	Glycogen operon protein	GH13	3	1	2	2		
3.2.1.54	Cyclomaltodextrinase	GH13/57	1	1				
3.2.1.14	Chitinase	GH18/19/48					1	
3.2.1.45	Glucosylceramidase	GH3/30			1	1		
3.2.1.26	β-fructosidase	GH32			1	2	2	

3.2.1.18	Sialidase	GH33/34/83	2	2				
3.2.1.22	α-galactosidase	GH4/27/36/57/97	1		6	3	2	3
3.2.1.24	α-mannosidase	GH31/38/47/92				1		2
3.2.1.89	Endogalactanase	GH53			2			
2.4.1.25	4-α-glucanotransferase	GH13/57/77	1	1	2	2	2	
3.2.1.96	Endo-β-N-acetyl glucosinaminidase	GH18/73/85/111		3				
3.2.1.52	β-N-acetylglucosaminidase	GH20/84/85	1	2				
3.2.1.140	Lacto-N-biose phosphorylase	GH20/112	3	2				
3.2.1.52	β-N-acetyl hexosaminidase	GH3/20	5	1		1		
3.2.1.55	α-L-Arabinosidase/arabinofuranosidase	GH3/43/51	1	1	4	6	1	3
3.2.1.37	β-xylosidase	GH1/3/30/39/43/52/116			5	3		4
3.2.1.41	Pullulanase	GH13			2	2	4	
3.2.1.17	1,4-β-N-acetylmuramidase	GH25/108	2	1	2	1	1	
3.2.1.25	β-mannosidase	GH1/2/5			2			
3.2.1.177	α-xylosidase	GH31			1	1		1
3.2.1.8	Endo-1,4-β-xylanase	GH5/8/10/11/30/43/51	1	1	3	2		3
3.2.1.93	Trehalose-6-phosphate hydrolase	GH13					1	
3.2.1.86	6-phospho- β -glucosidase	GH1		1				1
3.2.1.51	α-L-fucosidase	GH29		1	1			1
3.2.1.50	α-N-acetylglucosaminidase	GH89	1					
3.2.1.58	Glucan-1,3-β-glucosidase	GH3/5/17			2	2	1	
3.2.1.97	Endo-α-N-acetylgalactosaminidase	GH109	1					
3.2.1.4	cellulase/endo-1,4-β-glucanase	GH5/8			2		2	
3.2.1.86	6-phospho-β-glucosidase	GH1/4		1				1
3.2.1.67	1,4-α-galacturonidase	GH28						1
3.2.1.11	Pectinesterase	GH28						1

2.4.1.20	Cellibiose phosphorylase	GH94			1			
2.4.1.25/3.2.1.33	Glycogen debranching Enzyme	GH13					2	
3.2.1.40	α-L-rhamnosidase	GH78/106			1			
3.2.1.75	Glucan-endo-1,6-β-glucosidase	GH5/30						
3.2.1.78	Mannon-endo-1,4-β mannosidase	GH2/5			1	1		
3.2.1.139	α-glucoronidase	GH4/67						1
3.2.1.16	Trehalose synthase	GH13					1	
3.2.1.111	α-L-fucosidase 1,3/4	GH29/95	1					
3.2.1.51	α-L-1,2-fucosidase	GH29/95	1					
NA	Unknown	GH5						1
NA	Unknown	GH13		3	1	1		
NA	Unknown	GH23	2	1	2	2		1
NA	Unknown	GH30						3
NA	Unknown	GH43	1	1	2			1
NA	Unknown	GH101		1				
NA	Unknown	GH125			1			
NA	Unknown	GH127			2	1		
NA	Unknown	GH129	1	1				
NA	Unknown	GH120				1		
NA	Unknown	GH32						
NA	Unknown	GH95		1				
NA	Unknown	NC*			1			
NA	Unknown	GH65						1
NA	Unknown	GH2			1	1		
NA	Unknown	GH123	1	1				
NA	Unknown	GH3	1	1				

NA	Unknown	GH110	1	1				
NA	Unknown	GH33		1				
NA	Unknown	GH1				1		
NA	Unknown	GH105						1
NA	Unknown	GH115						2
NA	Unknown	GH25			1		2	1
Total			43	39	87	58	31	48

NC* indicates Not Characterised

Table 1.4.3 Bifdobacterial glycosyl hydrolases

		Cazy family	*B. animalis* subsp. *lactis* ATCC 27673	*B.animalis* subsp. *lactis* BLC1	*B.animalis* subsp. *lactis* Bi-07	*B. animalis* subsp. *lactis* Bl-04	*B. animalis* subsp. *lactis* Bl12	*B. animalis* subsp. *lactis* CNCM I-2494
EC Number	Enzyme name							
3.2.1.21	β-glucosidase/β-D-glucoside glucohydrolase	GH1/3	2	4	3	3	4	3
3.2.1.23	β-galactosidase	GH1/2/35/42	2	4	5	4	4	4
3.2.1.10	Oligo-1,6-glucosidase	GH13	1	1	1	1	1	1
3.2.1.20	α-glucosidase	GH4/13/31/63/97	1	1	1	1	1	1
2.4.1.7	Sucrose phosphorylase	GH13/57/119	2	2	2	1	2	1
3.2.1.1	α-amylase/neopullalanase	GH13/57	1	3	2	1	3	1
2.4.1.18	α-1,4-glucan branching enzyme	GH13/57	1	1	1	1	1	2
3.2.1.68	Glycogen operon protein	GH13	2	1		2	1	
3.2.1.45	Glucosylceramidase	GH3/30			2			2
3.2.1.26	β-fructosidase	GH32						1

3.2.1.22	α-galactosidase	GH4/27/36/57/97	3	3	3	3	3	3
2.4.1.25	4-α-glucanotransferase	GH13/57/77	2	2	2	2	2	2
3.2.1.52	β-N-acetylglucosaminidase	GH20/84/85						2
3.2.1.52	β-N-acetyl hexosaminidase	GH3/20	1		1	1		1
3.2.1.55	α-L-Arabinosidase/arabinofuranosidase	GH3/43/51	2	2	1	2	2	1
3.2.1.37	β -xylosidase	GH1/3/30/39/43/52/116	1	1	3	1	1	3
3.2.1.41	Pullulanase	GH13	1	2	4	2	1	4
3.2.1.17	1,4-β-N-acetylmuramidase	GH25/108	1	1	1	2	2	
3.2.1.25	β-mannosidase	GH1/2/5	2	1		2	1	
3.2.1.8	Endo-1,4-β-xylanase	GH5/8/10/11/30/43/51	1		1			
3.2.1.93	Trehalose-6-phosphate hydrolase	GH13	1					
3.2.1.4	Cellulase/endo-1,4-β-glucanase	GH5/8	1					
2.4.1.20	Cellibiose phosphorylase	GH94	1		1			1
2.4.1.25/3.2.1.33	Glycogen debranching Enzyme	GH13		1	2		1	2
3.2.1.75	Glucan-endo-1,6-β-glucosidase	GH5/30	1	1		1	1	
3.2.1.78	Mannon-endo-1,4-β-mannosidase	GH2/5						1
1.8.1.9	Unknown	NC	1	1		1	1	
3.2.1.16	Trehalose synthase	GH13			1	1		
NA	Unknown	GH13	1			1		1
NA	Unknown	GH23	2	1	2	2	1	1
NA	Unknown	GH30		1		1	1	
NA	Unknown	GH127	1	1	1	1	1	1
NA	Unknown	GH94		1		1		
NA	Unknown	GH25		2			1	1
Total			35	38	40	38	36	40

Table 1.4.4 Bifdobacterial glycosyl hydrolases

		Cazy family	*B. longum* subsp. *longum* BBMN68	*B. longum* subsp. *longum* F8	*B. longum* subsp. *longum* JCM 1217	*B. longum* subsp. *longum* JDM301	*B. longum* subsp. *longum* KACC 91563
EC Number	Enzyme name						
3.2.1.21	Glucosidase/β-D-glucoside glucohydrolase	GH1/3	2	3	2	3	3
3.2.1.23	β-galactosidase	GH1/2/35/42	3	4	3	7	4
3.2.1.10	Oligo-1,6-glucosidase	GH13		2	2		2
3.2.1.20	α-glucosidase	GH4/13/31/63/97	3	3	2		2
2.4.1.7	Sucrose phosphorylase	GH13/57/119	1	1	2	1	1
3.2.1.1	α-amylase/neopullalanase	GH13/57	4		1	6	3
2.4.1.18	α-1,4-glucan branching enzyme	GH13/57	1	1	1	1	1
3.2.1.54	Cyclomaltodextrinase	GH13/57		1			

3.2.1.26	β-fructosidase	GH32		1	1	2	
3.2.1.22	α-galactosidase	GH4/27/36/57/97	3	2	3	3	3
3.2.1.24	α-mannosidase	GH31/38/47/92	2	2		3	
2.4.1.25	4-α-glucanotransferase	GH13/57/77	2	2	2	2	2
3.2.1.52	β-N-acetylglucosaminidase	GH20/84/85	1	1		2	
3.2.1.140	lacto-N-biose phosphorylase	GH20/112	1		1		1
3.2.1.52	β -N-acetyl hexosaminidase	GH3/20		1	1		2
3.2.1.55	A-L-Arabinosidase/arabinofuranosidase	GH3/43/51	3	4	7	5	5
3.2.1.37	β-xylosidase	GH1/3/30/39/43/52/116	3	5		4	4
3.2.1.41	Pullulanase	GH13	2		1		
3.2.1.17	1,4-β-N-acetylmuramidase	GH25/108	1	1			
3.2.1.177	α-xylosidase	GH31			1	1	
3.2.1.8	Endo-1,4-β-xylanase	GH5/8/10/11/30/43/51			2	2	1
3.2.1.93	Trehalose-6-phosphate hydrolase	GH13	1	1			
3.2.1.68	Isoamylase	GH13		1			
3.2.1.89	Arabinogalactan endo-1,4-β-galactosidase	GH53		1	1	1	1
3.2.1.50	α-N-acetylglucosaminidase	GH89	1				
3.2.1.58	Glucan-1,3-β-glucosidase	GH3/5/17			1	1	1
3.2.1.81	β-agarase	GH16/50/86/118				1	
3.2.1.97	Endo-α-N-acetylgalactosaminidase	GH109			1		1
3.2.1.88	β-arabinosidase	GH		1			
3.2.1.4	Cellulase/endo-1,4-β-glucanase	GH5/8	1	1			
3.2.1.31	β-glucuronidase	GH1/2/79			1	1	
2.4.1.25/3.2.1.33	Glycogen debranching Enzyme	GH13		1	2	3	2
3.2.1.16	Trehalose synthase	GH13		1			1
3.2.1.111	α-L-fucosidase 1,3/4	GH29/95				1	

NA	Unknown	GH5	1	1		1	
NA	Unknown	GH13	4	1	4	1	3
NA	Unknown	GH23	2	3	2		3
NA	Unknown	GH30	2		2	1	1
NA	Unknown	GH43	8	2	8	3	3
NA	Unknown	GH101	1				
NA	Unknown	GH112		1		1	
NA	Unknown	GH121	1	1			1
NA	Unknown	GH125	1	1		1	
NA	Unknown	GH127	2	1	1	1	2
NA	Unknown	GH129		1	1	1	1
NA	Unknown	GH120		1			
NA	Unknown	GH32					1
NA	Unknown	GH95				1	
NA	Unknown	NC			1	1	
NA	Unknown	GH3	2		1		
NA	Unknown	GH25				2	
NA	Unknown	GH53	1				
Total			60	54	58	64	55

NC* indicates Not Characterised

1.4.4 *BIFIDOBACTERIAL α-GALACTOSIDASES*

Several putative and characterised α-galactosidases have been identified in bifidobacteria (71 in total based on Tables 1.4-1.8), of which most are very similar (belonging to GH family 36) and involved in the same carbohydrate metabolic pathway. This commonly occurring GH family 36 bifidobacterial α-galactosidase, is involved in the breakdown of raffinose family oligosaccharides (RFOs) (Martínez-Villaluenga & Gómez, 2007; O'Connell *et al.*, 2013b; Trojanova *et al.*, 2006; Turroni *et al.*, 2012c). This common α-galactosidase has been characterised in several bifidobacteria; for example the α-galactosidase produced by *B. breve* 203, was demonstrated to be capable of hydrolysing the α-(1→6) linkages of melibiose (Zhao *et al.*, 2008). More recently (and presented in this thesis) an α-galactosidase from *B. breve* UCC2003 (and nearly identical to that characterised for *B. breve* 203) was characterised, which was observed to hydrolyse the α-(1→6) bonds in stachyose and raffinose to produce sucrose and galactose, and cleave melibiose to generate glucose and galactose. This enzyme also showed hydrolytic activity against synthetic α-(1→4) and α-(1→3)-galactobiose therefore exhibiting a rather broad substrate range (O'Connell *et al.*, 2013b).

Other examples of α-galactosidases characterised in bifidobacteria are those found in *B. longum* (Garro *et al*., 1994; Hirayama *et al*., 2012), *B.adolescentis* DSM 20083 (Leder *et al.*, 1999; Van Laere *et al.*, 1999) and *B. bifidum* NCIMB 41171 (Goulas *et al.*, 2009). Interestingly, an α-galactosidase belonging to GH family 110 was identified in *B. bifidum* JCM 1254, that was shown to be capable of hydrolysing α-(1→3)-linked galactose bonds present in branched blood group B antigen, while also hydrolysing group B human salivary mucin and erythrocytes (Wakinaka *et al.*, 2013). This study would suggest that the α-galactosiase enzyme RafA presented in

this study may also possess the ability to degrade these very complex carbohydrates through cleavage of specific α-galactose moieties.

1.4.5 *BIFIDOBACTERIAL β-GALACTOSIDASES*

A significant number of putative and characterised β-galactosidases have been identified in bifidobacteria (128 in total based on Tables 1.4-1.8), of which many are very similar (belonging to GH family 2 and 42, with several also classified as GH families 1 and 35). β-galactosidases represent the best studied bifidobacterial GHs which have been characterised in a range of bifidobacteria, such as *B. angulatum* (Rabiu *et al*., 2001), *B. adolescentis* (Hinz *et al*., 2004), *B. bifidum* (Hattori & Taylor, 2009; Jorgensen *et al*., 2001; Moller *et al*., 2001), *B. longum* (Hsu *et al*., 2007; Li *et al*., 2014; Rossi *et al*., 2000), *B. longum subsp. infantis* (Hung & Lee, 2002; Moller *et al*., 2001; Viborg *et al*., 2013; Yoshida *et al*., 2012), *B. pseudolongum* (Rabiu *et al*., 2001), and *B. breve* (O'Connell Motherway *et al*., 2013; Yi *et al*., 2011), to name but a few.

The identified β-galactosidases identified in bifidobacteria have been shown to hydrolyse different substrates. For example, three β-galactosidases from *B. bifidum* were shown to hydrolyse the β-(1→4)-galactosidic bonds in lactose (Jorgensen *et al*., 2001), while a similar substrate specificity was observed for a β-galactosidase from *B. adolescentis* (Hinz *et al*., 2004). More recently a distinct hydrolytic ability was observed in the case of three β-galactosidases from *B. longum* subsp. *infantis* ATCC 15697, namely Bga42A, Bga42B, and Bga42C (Viborg *et al*., 2013). Bga42A was observed to prefer the β-(1→3)-galactosidic linkage found in certain HMOs, as well as other β-(1→3)- and β-(1→6)-galactosides attached to a glucose or galactose. In contrast, Bga42B very efficiently hydrolyses galactosyllactose, galactobiose and

galactotriose, which possess β-(1→4)-galactosidic linkages, while the substrate specificity of Bga42C resembled that of Bga42B, though with a lower specific activity.

Not only have several β-galactosidases been shown to hydrolyse glycosidic bonds of different substrates, they have also been demonstrated perform glycosyl transfer reactions to synthesise GOS from lactose (Rabiu *et al.*, 2001). The latter reactions require unnatural conditions: at a lactose concentration ranging from 30 % to 40 % (wt/wt), under these conditions GOS yields of 24.7 to 47.6% can be achieved (Hsu *et al.*, 2007; Rabiu *et al.*, 2001).

1.4.6 BIFIDOBACTERIAL α-GLUCOSIDASES

Quite a high number of α-glucosidases (50 in total based on the data presented in Tables 1.4-1.8) (most of which belong to GH families 13 or 31, although members of GH families 4, 63 and 97 have also been identified) have been either putatively identified or characterised in the genomes of various bifidobacterial species and strains. In *B. adolescentis* DSM 20083 two α-glucosidase-encoding genes (*aglA* and *aglB*) belonging to GH family 13 were characterised. Recombinant AglA demonstrated specificity for α-1,6 bonds, while AglB demonstrated specificity for α-(1→2/ α-1→3/α-1→4/α→1,6) bonds, although with a preference for α-(1→4) (van den Broek *et al.*, 2003).

Similarly, two α-glucosidases-encoding genes, *agl1* and *agl2*, were identified in *B. breve* UCC2003 (Pokusaeva *et al.*, 2009). The corresponding products, Agl1 and Agl2, were shown to exhibit hydrolytic activity towards the α-glucosidic linkages present in panose, isomaltose, isomaltotriose, and in four sucrose isomers, namely palatinose, trehalulose, turanose, and maltulose, while also hydrolysing such bonds

in trehalose and, to a lesser extent, nigerose. In this instance the substrates of preference for both enzymes were panose, isomaltose, and trehalulose, representing α-(1→6) and α-(1→1) glucosidic linkages. Recently, a third α-glucosidase encoded by *B. breve* UCC2003 was characterised (presented in this thesis), which was shown to hydrolyse melezitose into glucose and fructose, thereby demonstrating that this protein has both α-(1→2) and α-(1→3)-glucosyl hydrolase activities. This protein was furthermore observed to hydrolyse sucrose and turanose to glucose and fructose (O'Connell *et al.*, 2013b). Other α-glucosidases are predicted to be encoded by *B. breve* UCC2003, indicating that this strain encodes additional specificities that allow this strain to access other carbohydrate substrates with α-glucosidic linkages.

1.5 DE NOVO MICROBIAL BIOSYNTHESIS OF FATTY ACIDS IN ESCHERICHIA COLI

In this thesis we have studied the biosynthesis of certain unusual fatty acids. First, we will in broad terms discuss biosynthesis of fatty acids by Gram negative bacteria (*E. coli*), complemented with some additional information on this metabolic process in a high GC Gram positive bacteria (*Streptomyces coelicolor*). Fatty acids are generally soluble in organic compounds and are found in tissues of plants, animals and microorganisms, particularly when they are linked to glycerol molecules in the form of lipids (Christie, 2003). The typical molecular structure of a fatty acid is a straight, even-numbered chain of carbons terminated by a carboxyl group at one end. These fatty acids can take one of two forms, either being saturated when the fatty acid contains no double bonds, or being unsaturated implying that the fatty acid contains one or more double bonds in either *cis* or *trans* conformation (NB. in the *cis* conformation, hydrogen atoms are on the same side of the double bond, whereas in the *trans* conformation, they are on opposite sides) (Christie, 2003).

De novo fatty acid biosynthesis represents a crucial metabolic pathway in all living organisms. Fatty acid synthesis can be achieved via so-called type I or type II pathways. The main difference separating these two biosynthetic pathways is the way in which the fatty acid biosynthetic machinery is encoded. In the case of the type II system, which exists primarily in bacteria and plants, the enzymatic steps required for fatty acid biosynthesis are represented by individual proteins that are expressed by separate genes (White *et al.*, 2005). In contrast, type I systems, which are mainly found in eukaryotes, carry all enzymatic activities required for fatty acid biosynthesis on one or two large polypeptide chains (Schweizer & Hofmann, 2004).

Fatty acid biosynthesis has been studied in considerable detail in *E. coli* (Chan & Vogel, 2010; Rock & Cronan, 1996). For *de novo* (saturated) fatty acid production the fatty acid biosynthesis process begins with the carboxylation of acetyl-CoA into malonyl-CoA using acetyl-CoA carboxylase (Polakis *et al.*, 1974, Fig. 1.2 A, step 1). Malonyl-CoA-ACP transacylase, specified by the FabD protein, transfers malonyl-CoA to ACP (Acyl carrier protein), which is a cofactor protein that covalently binds all fatty acyl intermediates (Ruch & Vagelos, 1973) (Fig. 1.2 A, step 2). The formed malonyl-ACP then combines with an acetyl-CoA group by means of a condensation reaction that is catalysed by β-oxoacyl synthase III, also designated FabH, to form β-oxobutyryl-ACP (Tsay *et al.*, 1992) (Fig. 1.2 A, step 3). Following this step, the β-oxoacyl-ACP is reduced by an NADPH-dependent β-oxoacyl reductase, designated FabG, into β-hydroxyacyl-ACP (Alberts *et al.*, 1964) (Fig. 1.2 A, step 4). The hydroxy group is removed by one of two β-hydroxyacyl dehydratases, represented by FabA (Birge *et al.*, 1967; Kass *et al.*, 1967; Leon & Bloch, 1967) or FabZ (Mohan *et al.*, 1994), which converts the chain into a *trans*-2-enoyl group (Fig. 1.2 A, step 5). This leads to the reduction of the double bond in an NADH-dependent

reaction by an enoyl-reductase, designated FabI (Weeks & Wakil, 1968) (Fig. 1.2 A, step 6). Acyl-ACP bound by a saturated acyl chain extended by two carbon units is produced at this point. This cycle may also be initiated by FabB or FabF (Fig. 1.2 A, step 7).

De novo production of unsaturated fatty acids (Chan & Vogel, 2010) follows in many ways the same biosynthetic pathway used for saturated fatty acid production as described above. The first step in the production of these acyl chains is facilitated by the *fabA* gene product, a β-hydroxyacyl dehydratase, which catalyses both a dehydration as well as an isomerisation reaction. FabA converts β-hydroxydecanoyl-ACP into *trans*-2-decenoyl-ACP (Fig. 1.2 B, step 1) via a dehydration reaction and subsequently isomerises this fatty acyl intermediate trans-2-decenoyl-ACP to *cis*-3-decenoyl-ACP (Fig. 1.2 B, step 2) (Birge *et al.*, 1967; Kass *et al.*, 1967; Leon & Bloch, 1967). Subsequently, the *cis*-3-decenoyl group undergoes a condensation reaction with another malonyl-ACP group, yielding *cis*-5-dodecenoyl-ACP (Fig. 1.2 B, step 3), where the two added carbon groups are introduced closest to the carbonyl carbon. This reaction is catalysed by the second essential enzyme for unsaturated fatty acid synthesis, β-oxoacyl synthase I or FabB (Garwin *et al.*, 1980a; Garwin *et al.*, 1980b). This is then followed by the same elongation steps as for saturated fatty acid synthesis (Fig. 1.2 B, step 4 and 1.2A steps 1-7).

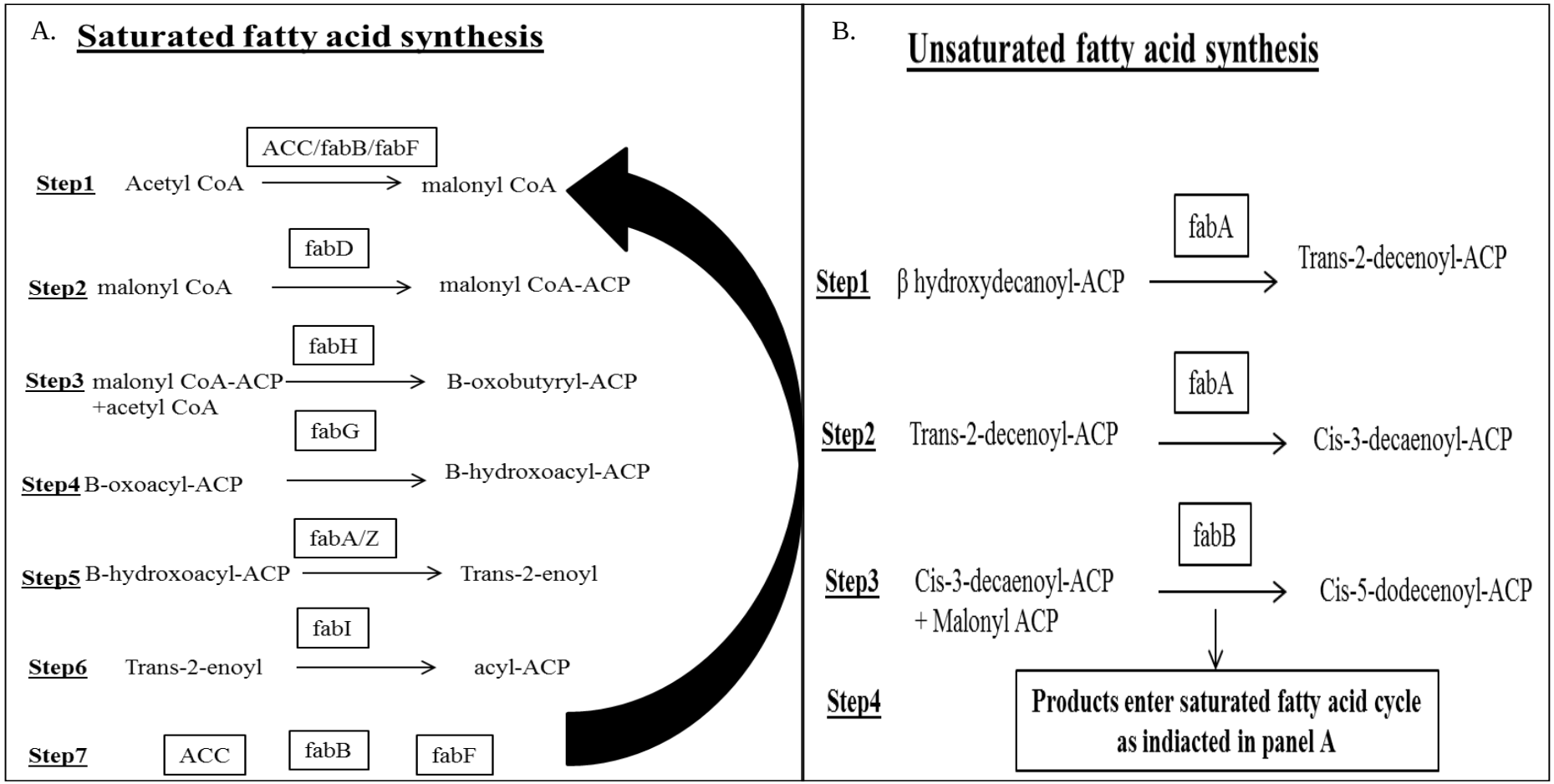

Figure 1.2 *De novo* microbial biosynthesis of fatty acids in *E. coli* of (Panel A) saturated fatty acids (ACC represents acetyl CoA carboxylase) and (Panel B) unsaturated fatty acids (adapted from (Chan & Vogel, 2010).

1.5.1 DE NOVO MICROBIAL BIOSYNTHESIS OF FATTY ACIDS IN STREPTOMYCETES

Although there is a paucity of information related to fatty acid biosynthesis in bifidobacteria, some relevant information is available from the Gram positive, high GC content *Streptomyces.* It was originally speculated that fatty acid synthesis in *Streptomyces* functioned as a type I system which is mainly present in yeast and mammalian cells (Revill *et al.*, 1995; Rossi & Corcoran, 1973). However, it was later realised that this was incorrect when a type II system for fatty acid biosynthesis was discovered in *Streptomyces* (Revill *et al.*, 1995; Summers *et al.*, 1995). In contrast to *E. coli* which produces mainly straight chain fatty acids and functions specifically with acetyl-CoA (Tsay *et al.*, 1992), species of *Streptomyces* most often produce branched fatty acids and can synthesise fatty acids employing other short chain fatty acids, such as isobutyryl-CoA and methylbutyryl-CoA, as well as acetyl-CoA, as feeding materials for this biosynthetic pathway (Han *et al.*, 1998).

The *acpP* gene product of *Streptomyces coelicolor* has been proven to function as an acyl carrier protein due to its ability to compensate for an *E. coli* mutant deficient in fatty acid synthesis (Revill *et al.*, 1996). Similarly, a gene which encodes FabD, specifying the malonyl-CoA-ACP transacylase, was identified in *Streptomyces glaucescens*, and shown to be able to complement an *E. coli* mutant carrying a dysfunctional *fabD* (Summers *et al.*, 1995). This *S. glaucescens fabD* gene was also shown to be active in the presence of acyl carrier proteins from *E. coli* and *S. glaucescens* (Florova *et al.*, 2002). The β-oxoacyl synthase III enzyme or FabH has also been identified in *Streptomycetes* through complementation of an *S. coelicolor* mutant carrying a mutated *fabH* gene. Interestingly, the presence of *fabH* from *E. coli* in this *S. coelicolor* mutant converted this complemented strain to the production

of mainly (86 %) straight chain fatty acids as opposed to the normal branched chain fatty acids (Li *et al.*, 2005). Fatty acid synthesis in *E. coli* and *S. coelicolor* both show similarity in their mechanisms and therefore one would speculate that these two models are an apt example of how fatty acid biosynthesis would occur in bifidobacteria.

1.6 CONJUGATED LINOLEIC ACID

In this thesis the focus is specifically on linoleic acid (LA) which cannot be synthesised by humans and is a polyunsaturated fatty acid (PUFA) that is present in animals, plants and microorganisms (Gill & Valivety, 1997). Conjugated linoleic acid (CLA), first identified over 25 years ago (Ha *et al.*, 1987), is the collective term used to describe a group of PUFA that exist as positional and geometric stereoisomers of octadecadienoic acid. Unsaturated fatty acids are divided into mono-unsaturated fatty acids (MUFA), i.e. fatty acids that contain just a single double bond, and PUFA, which possess more than one double bond (Christie, 2003). Many of the fatty acids can be synthesised by humans; however, there is one particular group of essential PUFA, which the human body cannot produce, and which include omega-3 and omega-6 fatty acids. Examples are LA, an omega-6 fatty acid essential for building the water barrier in the skin (Zheng *et al.*, 2011), and α-linolenic acid, an omega-3 fatty acid and a precursor for eicosapentaenoic acid (EPA) and docosahexaenoic acid (DHA), both representing fatty acids that exhibit a range of health benefits such as anti-inflammatory activity (for a review see (Wall *et al.*, 2010)). Twenty eight different CLA isomers have so far been identified, of which the most abundant is the *cis*-9 *trans*-11 (c9, t11) isomer, representing approximately 80 % of total CLA in food products. CLA is found naturally in ruminant food products, for example lamb, beef and dairy, due to the process of

biohydrogenation of LA by specific bacteria in the rumen (Bhattacharya *et al.*, 2006).

CLA has attracted a considerable amount of attention due to its purported health benefits. Interest in CLA grew when Pariza and his colleagues discovered the anti-carcinogenic properties of both raw and grilled ground beef, where active compounds were identified as four isomers of LA (Pariza, 1997). These findings were extended when it was shown that certain isomers of CLA reduce carcinogenesis and body fat (Ogawa *et al.*, 2005). Interest in CLA as a weight loss treatment has intensified in recent years, since it was first discovered that CLA ingestion can modulate body fat composition (Park *et al.*, 1997b). In a recent study the supplementation of CLA over a 12 week period in overweight obese Chinese subjects resulted in lower obesity indices (Chen *et al.*, 2012).

Prebiotics compounds have been shown to not only stimulate the increased presence of probiotic bacteria, but have also been linked to modulating the amount of conjugated linoleic acid produced by certain commensal bacteria (Druart *et al.*, 2013). *In vitro* experiments have shown that isolated human gut bacteria have the ability to convert polyunsaturated fatty acids such as LA into conjugated PUFA such as CLA. It has been observed that feeding mice a diet that is high in LA coupled with supplementation with prebiotic carbohydrates increases the *in vivo* production of CLA by the murine gut microbiota (Druart *et al.*, 2013).

1.6.1 MICROBIAL BIOHYDROGENATION

PUFAs are produced in the rumen by certain bacteria in ruminant animals. These ruminant animals are fed forage, which is a good source of LA and linolenic acid, while seed oils used to concentrate feed are also high in linoleic acid. It has been

hypothesised that following ingestion these dietary lipids undergo hydrolysis of the ester linkages by microbial lipases, this being a prerequisite for the second step, i.e. the biohydrogenation of fatty acids. Fatty acids are then released into the rumen, adsorbed on to feed particles and hydrogenated or integrated into bacterial lipids (Harfoot & Hazlewood, 1997; Jenkins *et al.*, 2008). The process of biohydrogenation involves conversion of dietary LA to CLA (c9, t11) by a bacterial enzyme known as linoleic acid isomerase (LAI). CLA can then be further reduced via a reductase to form vaccenic acid, which is then thought to be reduced to stearic acid (Fig. 1.3). Endogenous synthesis also occurs simultaneously in the animal's own tissues, the reduction of CLA to vaccenic acid is believed to be the rate limiting step in the reduction of CLA to stearic acid, which causes the build-up of vaccenic acid in the tissues where endogenous synthesis occurs via Δ9 desaturase enzyme (Griinari *et al.*, 2000) (Fig. 1.3). *Butyrivibrio fibrisolvens* was the first *bacterium* shown to perform the process of biohydrogenation of fatty acids (Polan *et al.*, 1964). CLA was shown to be produced as an intermediate of this process (Kepler *et al.*, 1966).

In recent years a number of other bacteria belonging to the genera *Propionibacterium, Lactobacillus* and *Bifidobacterium* have been reported to convert LA into CLA when grown in either synthetic media or milk (Barrett *et al.*, 2007). The biohydrogenation process is believed to be similar to that performed by ruminant bacteria (Lin *et al.*, 2002), although CLA may undergo further isomerisation due to the fact that the resulting CLA product is not reduced or adsorbed (Coakley *et al.*, 2003). It has been hypothesised that the process of CLA production by bifidobacteria may contribute to the positive effects of these bacteria in the GIT (Coakley *et al.*, 2006).

It has also been hypothesised that the conversion of free LA to CLA may function as a detoxification mechanism in bacteria as free LA has been shown to inhibit growth of CLA-producing bacteria (Jiang *et al.*, 1998). The bent structure of LA may interfere with the bacterial membrane and when free LA is conjugated the resulting fatty acid strain becomes straight chain fatty acid, thereby decreasing the fluidity of the membrane, which may also have a positive effect on stress resistance (Weber *et al.*, 1994).

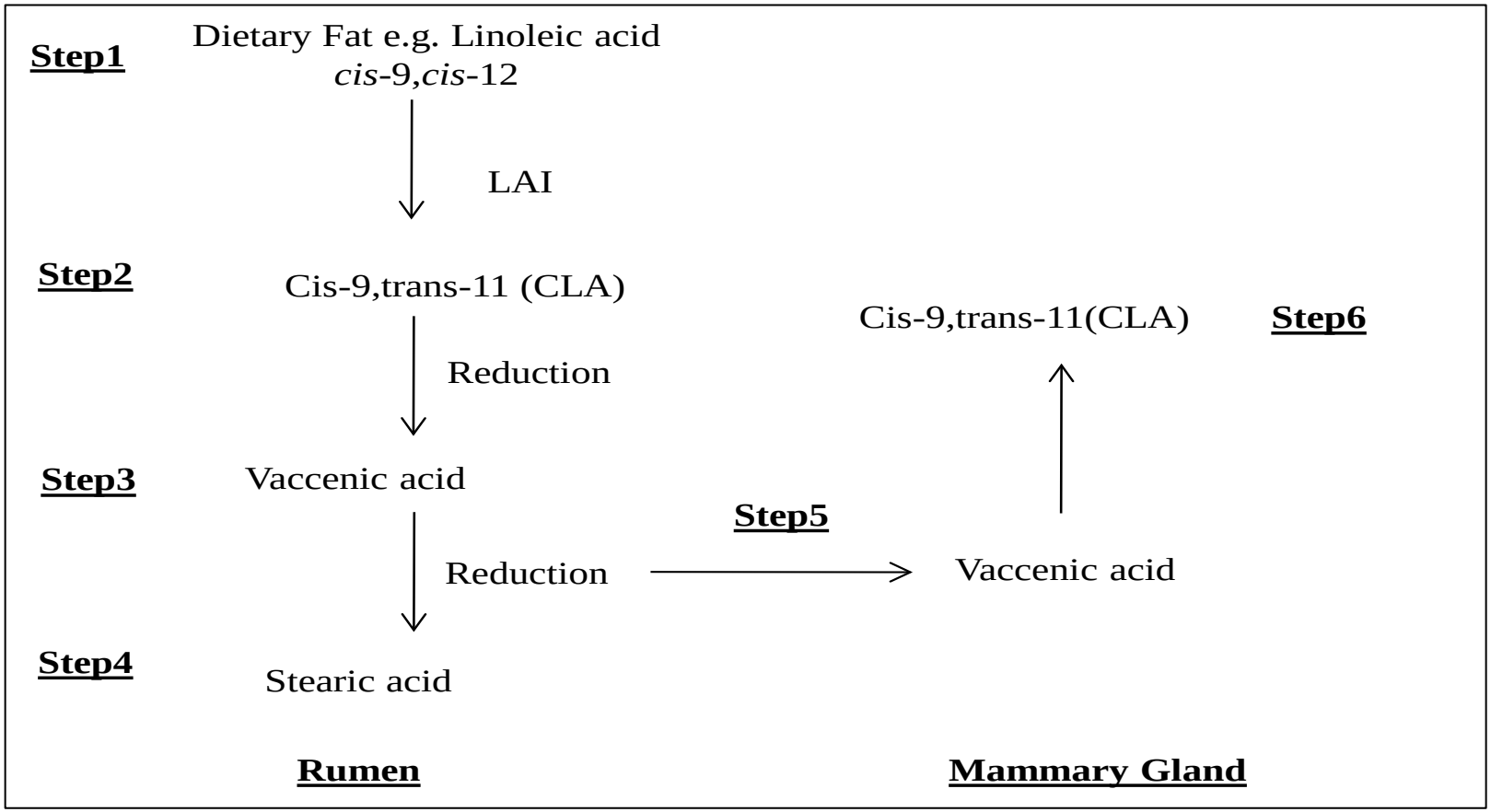

Figure 1.3 Microbial biohydrogenation (adapted from (Bauman *et al.*, 2000).

1.6.2 LINOLEIC ACID ISOMERASE

The bacterial enzyme that is known to be involved in the conversion of LA to c9, t11 CLA is the linoleic acid isomerase (LAI; EC 5.3.1.5). This enzyme was first isolated from *Butyrivibrio fibrisolvens* (Kepler *et al.*, 1966; Kepler *et al.*, 1970) and it was concluded from these studies that this isomerase is membrane bound, that it demonstrates a substrate requirement for *cis*-9,*cis*-12 diene systems and a free carboxyl group (as present in LA) (Kepler *et al.*, 1970). The LAI from *Propionibacterium acnes* which isomerises LA to t10, c12 CLA has been expressed in bacteria, yeast and tobacco seeds (Hornung *et al.*, 2005; Rosberg-Cody *et al.*, 2007). A gene encoding the LAI from *P. acnes* was cloned and overexpressed in *Lactococcus lactis* and *E. coli*, resulting in between 30-50 % conversion rates of LA to t10, c12 CLA (Rosberg-Cody *et al.*, 2007).

More recently the LAI enzymes from *P. acnes, Lb. reuteri* and *Lb. acidophilus* have also been expressed in *E. coli* (Luo *et al.*, 2013). In this study they achieved purification of an active protein which has proven difficult previously to purify and subsequently had the ability to test different parameters of the enzyme in order to achieve optimal activity (Luo *et al.*, 2013). It has been shown that lactobacilli expressing LAI from *P. acnes* can modulate the fatty acid composition of host adipose tissue in mice (Rosberg-Cody *et al.*, 2011b). A recombinant *Lactobacillus paracasei* NFBC 338 derivative expressing the LAI from *P. acnes* was tested in order to see if it had an effect on fatty acid composition of different mouse tissues. The result showed a 4-fold increase in t10, c12 CLA in adipose tissue of mice as compared to those which received the non-CLA producing strain. The liver of the mice which received the recombinant strain also showed a 2.5-fold increase in the level of t10, c12 CLA as compared to the control strain.

1.6.3 CLA PRODUCTION BY BIFIDOBACTERIA

In a study performed by (Coakley *et al.*, 2003) the ability to convert LA to CLA by *Lactobacillus, Lactococcus, Pediococcus* and *Bifidobacterium* was assessed. The *B. breve* strain NCFB 2258 was shown to be the most efficient in converting free LA to CLA, as it was demonstrated to convert 65 % of linoleic acid to the c9, t11 CLA isomer. In a subsequent study the anti-proliferative characteristics of c9, t11 and t9, t11 obtained from *B. breve* NCFB 2258 CLA production was tested on two colon cancer cell lines resulting in both showing anti-proliferative affect with t9, t11 being the most potent (Coakley *et al.*, 2006). *B. breve* DPC 6330, an isolate from a faecal sample from an elderly patient, showed high CLA-producing ability among the isolates tested (Barrett *et al.*, 2007). More recently 17 bifidobacterial strains were tested for LA to CLA bioconversion and of the 17 strains 11 were capable of LA to CLA bioconversion (O'Connell *et al.*, 2013a).

1.6.4 MYOSIN CROSS REACTIVE ANTIGEN

Myosin cross reactive (MCRA-like) proteins show over 50 % sequence similarity to the LAI from *Lb. acidophilus* and *Lb. reuteri* PYR8 (Rosson *et al.*, 2001). The first MCRA protein (MW ~67 kDa) was identified in *Streptococcus pyogenes*, while screening for antigens recognised by acute rheumatic fever sera (Kil *et al.*, 1994). The MCRA protein of *S. pyogenes* M49 was recently shown to represent a flavin adenine dinucleotide (FAD) enzyme which acts as a fatty acid hydratase (Volkov *et al.*, 2010), catalysing the conversion of oleic acid into 10-hydroxystearic acid (Joo *et al.*, 2012).

The MCRA-like protein from *B. breve* has recently been characterised as an FAD-dependent fatty acid hydratase with a role in stress protection (Rosberg-Cody *et al.*,

2011a). The *MCRA* gene from *B. breve* NCFB 2258 was expressed in *L. lactis* and *Corynebacterium glutamicum*, and the recombinant protein was then tested for enzymatic activity using the fatty acid substrates palmitoleic, oleic and LA. The bifidobacterial MCRA was shown to catalyse the conversion of these fatty acids to 10-hydroxy fatty acids. Other MCRA proteins have been characterised recently in hosts such as *Lactobacillus rhamnosus* LGG, *Lb. plantarum* ST-III, *Lb. acidophilus* NCFM and *B. animalis* subsp. *lactis* BB-12. These MCRA proteins were expressed in *E. coli*, and in all four cases the resulting recombinant *E. coli* strains catalysed the conversion of LA and oleic acid to 10-hydroxy derivatives (Yang *et al.*, 2012). In both cases purified proteins were used.

As mentioned previously MCRA-like proteins show over 50 % sequence similarity to the LAI from *Lb. acidophilus* and *Lb. reuteri* PYR8 (Rosson *et al.*, 2001). The microbial production of CLA was elucidated using washed cells of *Lb. acidophilus* AKU 1137 (Ogawa *et al.*, 2001) and involves the production of hydroxy fatty acids as precursor to formation of CLA. This study speculated that hydroxyl fatty acid production is the first step in CLA formation. When such hydroxy fatty acids are isolated and introduced to the washed cells of *Lb. acidophilus* AKU 1137, conversion to CLA isomers was shown to take place (Ogawa *et al.*, 2001). It was therefore concluded that CLA formation by *Lb. acidophilus* AKU 1137 involves two steps: hydration of LA to 10-hydroxy-18:1, and the subsequent dehydration and isomerisation of these hydroxy fatty acids to the c9, t11 CLA and t9, t11 CLA isomers. Based on this context it was hypothesised that the MCRA plays an important role in CLA production.

However, more recently (and presented in this thesis) it was observed that when an insertion mutant was created in the MCRA-encoding gene of *B. breve* NCFB 2258,

no effect on CLA production was observed, demonstrating that the MCRA-encoding gene does not play a role in CLA production by *B. breve* (O'Connell *et al.*, 2013a). The $MCRA_{UCC2003}$ insertion in *B. breve* UCC2003 did however affect its hydratase activity, as its ability to produce 10-hydroxy fatty acids from oleic acid was severely reduced.

1.7 CONCLUSION

Discovery of novel, efficient and clinically proven prebiotics is an important field of product development for the food industry. Investigations on what carbohydrates can be utilised by bifidobacteria and the identification of the metabolic pathways involved will provide clues as to the identity of novel prebiotics. Through transglycosylation reactions it may also be possible to develop novel compounds utilising bifidobacterial GHs. Genome sequencing has allowed the identification of many putative GH-encoding genes in bifidobactreia, and much research is still to be done to fully explore and exploit these carbohydrate-active enzymes.

Bifidobacteria have the ability to produce CLA which has been associated with many health benefits such as anti-tumour and weight loss activities. Genome sequencing and putative characterisation of genes has provided further insights into the metabolic process of fatty acid metabolism of bifidobacteria. However, the precise mechansism of CLA production by bifidobacteria is still elusive and further research has to be performed in order to fully elucidate this mechanism. It is also important to develop genetic tools capable of genetically manipulating bifidobacteria such as *E. coli*/bifidobacterial shuttle vectors as bifidobacteria are notoriously difficult to manipulate and there is a paucity of genetic tools available for this purpose.

1.7.1 THESIS OUTLINE

Chapter I (above) is a general introduction of this thesis which covers carbohydrate metabolism and fatty acid metabolism specifically in bifidobacteria.

Chapter II describes the identification and characterisation of a bifidobacterial oleate hydratase and its role in stress tolerance.

Chapter III reports the identification and characterisation of two clusters of *B. breve* UCC2003 involved in raffinose-like sugars and melezitose utilisation.

Chapter IV deals with the regulation of the gene clusters involved in the utilisation of raffinose-like sugars and melezitose utilisation by *B. breve* UCC2003.

Chapter V aims to characterise the replication function of a novel bifidobacterial megaplasmid and attempts to create a shuttle vector.

1.7.2 REFERENCES

Alberts, A. W., Majerus, P. W., Talamo, B. & Vagelos, P. R. (1964). Acyl-Carrier Protein. II. Intermediary Reactions of Fatty Acid Synthesis. *Biochemistry* **3**, 1563-1571.

Almeida, C. C., Lorena, S. L. S., Pavan, C. R., Akasaka, H. M. I. & Mesquita, M. A. (2012). Beneficial Effects of Long-Term Consumption of a Probiotic Combination of *Lactobacillus casei Shirota* and *Bifidobacterium breve* Yakult May Persist After Suspension of Therapy in Lactose-Intolerant Patients. *Nutrition in Clinical Practice* **27**, 247-251.

Alvarez-Martin, P., O'Connell-Motherway, M., van Sinderen, D. & Mayo, B. (2007). Functional analysis of the pBC1 replicon from *Bifidobacterium catenulatum* L48. *Applied Microbiology and Biotechnology* **76**, 1395-1402.

Andersen, J., Barrangou, R., Hachem, M., Lahtinen, S., Goh, Y., Svensson, B. & Klaenhammer, T. (2013). Transcriptional analysis of oligosaccharide utilisation by *Bifidobacterium lactis* Bl-04. *BMC Genomics* **14**, 312.

Bäckhed, F., Ley, R. E., Sonnenburg, J. L., Peterson, D. A. & Gordon, J. I. (2005). Host-Bacterial Mutualism in the Human Intestine. *Science* **307**, 1915-1920.

Barrangou, R., Briczinski, E. P., Traeger, L. L. Loquasto, J.R., Richards, M., Horvath, P., Coûté-Monvoisin, A.C., Leyer, G., Rendulic, S., Steele, J.L., Broadbent, J.R., Oberg, T., Dudley, E.G., Schuster, S., Romero, D.A. & Roberts, R.F. (2009). Comparison of the complete genome sequences of *Bifidobacterium animalis* subsp. *lactis* DSM 10140 and Bl-04. *Journal of Bacteriology* **191**, 4144-4151.

Barrett, E., Ross, R. P., Fitzgerald, G. F. & Stanton, C. (2007). Rapid Screening Method for Analysing the Conjugated Linoleic Acid Production Capabilities of Bacterial Cultures. *Applied and Environmental Microbiology* **73**, 2333-2337.

Bauman, D. E., Baumgard, L. H., Corl, B. A. & Griinari, J. M. (2000). Biosynthesis of conjugated linoleic acid in ruminants. *Journal of Animal Science* **77**, 1-15.

Bhattacharya, A., Banu, J., Rahman, M., Causey, J. & Fernandes, G. (2006). Biological effects of conjugated linoleic acids in health and disease. *The Journal of Nutritional Biochemistry* **17**, 789-810.

Bibel, D. J. (1988). Elie Metchnikoff's bacillus of long life. *ASM News* **54**, 661-665.

Birge, C. H., Silbert, D. F. & Vagelos, P. R. (1967). A β-hydroxydecanoyl-ACP dehydrase specific for saturated fatty acid biosynthesis in *E. coli*. *Biochemical and Biophysical Research Communications* **29**, 808-814.

Bjorklund, M., Ouwehand, A. C., Forssten, S. D., Nikkila, J., Tiihonen, K., Rautonen, N. & Lahtinen, S. J. (2012). Gut microbiota of healthy elderly NSAID users is selectively modified with the administration of *Lactobacillus acidophilus* NCFM and lactitol. *Age* **34**, 987-999.

Blum, S. & Schiffrin, E. J. (2003). Intestinal microflora and homeostasis of the mucosal immune response: implications for probiotic bacteria? *Current Issues in Intestinal Microbiology* **4**, 53-60.

Bordoni, A., Amaretti, A., Leonardi, A., Boschetti, E., Danesi, F., Matteuzzi, D., Roncaglia, L., Raimondi, S. & Rossi, M. (2013). Cholesterol-lowering probiotics: *in vitro* selection and *in vivo* testing of bifidobacteria. *Applied Microbiology and Biotechnology* **97**, 8273-8281.

Bottacini, F., Dal Bello, F., Turroni, F., Milani, C., Duranti, S., Foroni, E., Viappiani, A., Strati, F., Mora, D., van Sinderen, D. & Ventura, M. (2011). Complete genome sequence of *Bifidobacterium animalis* subsp. *lactis* BLC1. *Journal of Bacteriology* **193**, 6387-6388.

Bottacini, F., Milani, C., Turroni, F., Sánchez, B., Foroni, E., Duranti, S., Serafini, F., Viappiani, A., Strati, F., Ferrarini, A., Delledonne, M., Henrissat, B., Coutinho, P., Fitzgerald, G.F., Margolles, A., van Sinderen D. & Ventura, M. (2012). *Bifidobacterium asteroides* PRL2011 genome analysis reveals clues for colonisation of the insect gut. *PLoS One* **7**, e44229.

Bourget, N., Simonet, J.-M. & Decaris, B. (1993). Analysis of the genome of the five *Bifidobacterium breve* strains: plasmid content, pulsed-field gel electrophoresis genome size estimation and rrn loci number. *FEMS Microbiology Letters* **110**, 11-20.

Bramhill, D. & Kornberg, A. (1988). Duplex opening by dnaA protein at novel sequences in initiation of replication at the origin of the *E. coli* chromosome. *Cell* **52**, 743-755.

Cantarel, B. L., Coutinho, P. M., Rancurel, C., Bernard, T., Lombard, V. & Henrissat, B. (2009). The Carbohydrate-Active EnZymes database (CAZy): an expert resource for Glycogenomics. *Nucleic Acids Research* **37**, D233-238.

Chan, D. I. & Vogel, H. J. (2010). Current understanding of fatty acid biosynthesis and the acyl carrier protein. *The Biochemical Journal* **430**, 1-19.

Chen, S.-C., Lin, Y.-H., Huang, H.-P., Hsu, W.-L., Houng, J.-Y. & Huang, C.-K. (2012). Effect of conjugated linoleic acid supplementation on weight loss and body fat composition in a Chinese population. *Nutrition* **28**, 559-565.

Chervaux, C., Grimaldi, C., Bolotin, A., Quinquis, B., Legrain-Raspaud, S., van Hylckama Vlieg, J. E., Denariaz, G. & Smokvina, T. (2011). Genome sequence of the probiotic strain *Bifidobacterium animalis* subsp. *lactis* CNCM I-2494. *Journal of Bacteriology* **193**, 5560-5561.

Christie, W. W. (2003). *Lipids: their structure and occurence.* Bridgewater, England: The Oily Press.

Coakley, M., Ross, R. P., Nordgren, M., Fitzgerald, G., Devery, R. & Stanton, C. (2003). Conjugated linoleic acid biosynthesis by human-derived *Bifidobacterium* species. *Journal of Applied Microbiology* **94**, 138-145.

Coakley, M., Johnson, M. C., McGrath, E., Rahman, S., Ross, R. P., Fitzgerald, G. F., Devery, R. & Stanton, C. (2006). Intestinal Bifidobacteria That Produce trans-9, trans-11 Conjugated Linoleic Acid: A Fatty Acid With Antiproliferative Activity Against Human Colon SW480 and HT-29 Cancer Cells. *Nutrition and Cancer* **56**, 95-102.

Collado, M. C., Bauerl, C. & Perez-Martinez, G. (2012). Defining microbiota for developing new probiotics. *Microbial Ecology in Health and Disease* **23**.

Corneau, N., Emond, E. & LaPointe, G. (2004). Molecular characterisation of three plasmids from *Bifidobacterium longum*. *Plasmid* **51**, 87-100.

Cousin, F. J., Foligné, B., Deutsch, S.-M., Massart, S., Parayre, S., Le Loir, Y., Boudry, G. & Jan, G. (2012). Assessment of the Probiotic Potential of a Dairy Product Fermented by *Propionibacterium freudenreichii* in Piglets. *Journal of Agricultural and Food Chemistry* **60**, 7917-7927.

Cronin, M., Knobel, M., O'Connell-Motherway, M., Fitzgerald, G. F. & van Sinderen, D. (2007). Molecular Dissection of a Bifidobacterial Replicon. *Applied and Environmental Microbiology* **73**, 7858-7866.

Cronin, C., Ventura , M., Fitzgerald, G.F. & van Sinderen, D. (2011). Progress in genomics, metabolism and biotechnology of bifidobacteria. *International Journal of Food Microbiology* **149,** 4-18.

Cummings, J. H. & Macfarlane, G. T. (1991). The control and consequences of bacterial fermentation in the human colon. *The Journal of Applied Bacteriology* **70**, 443-459.

Dahl, M. K., Degenkolb, J. & Hillen, W. (1994). Transcription of the xyl operon is controlled in *Bacillus subtilis* by tandem overlapping operators spaced by four base-pairs. *Journal of Molecular Biology* **243**, 413-424.

Davies, G. & Henrissat, B. (1995). Structures and mechanisms of glycosyl hydrolases. *Structure* **3**, 853-859.

Deutscher, J., Francke, C. & Postma, P. W. (2006). How phosphotransferase system-related protein phosphorylation regulates carbohydrate metabolism in bacteria. *Microbiology and Molecular Biology Reviews* **70**, 939-1031.

Druart, C., Neyrinck, A. M., Dewulf, E. M., De Backer, F.C., Possemiers, S., Van de Wiele, T., Moens, F., De Vuyst, L., Cani, P.D., Larondelle, Y. & Delzenne, N.M. (2013). Implication of fermentable carbohydrates targeting the gut microbiota on conjugated linoleic acid production in high-fat-fed mice. *British Journal of Nutrition* **110**, 998-1011.

Dubeau, M.-P., Poulin-Laprade, D., Ghinet, M. G. & Brzezinski, R. (2011). Properties of CsnR, the Transcriptional Repressor of the Chitosanase Gene, csnA, of *Streptomyces lividans*. *Journal of Bacteriology* **193**, 2441-2450.

Eckburg, P. B., Bik, E. M., Bernstein, C. N., Purdom, E., Dethlefsen, L., Sargent, M., Gill, S. R., Nelson, K. E. & Relman, D. A. (2005). Diversity of the Human Intestinal Microbial Flora. *Science* **308**, 1635-1638.

Fanning, S., Hall, L. J., Cronin, M. Zomer, A., MacSharry, J., Goulding, D., O'Connell Motherway, M., Shanahan, F., Nally, K., Dougan, G. & van Sinderen, D. (2012). Bifidobacterial surface-exopolysaccharide facilitates commensal-host interaction through immune modulation and pathogen protection. *Proceedings of the National Academy of Sciences* **109**, 2108-2113.

Flint, H. J., Scott, K. P., Duncan, S. H., Louis, P. & Forano, E. (2012). Microbial degradation of complex carbohydrates in the gut. *Gut Microbes* **3**, 289-306.

Florova, G., Kazanina, G. & Reynolds, K. A. (2002). Enzymes involved in fatty acid and polyketide biosynthesis in *Streptomyces glaucescens*: role of FabH and FabD and their acyl carrier protein specificity. *Biochemistry* **41**, 10462-10471.

Francavilla, R., Lionetti, E., Castellaneta, S., Ciruzzi, F., Indrio, F., Masciale, A., Fontana, C., La Rosa, M.M., Cavallo, L. & Francavilla, A. (2012). Randomised clinical trial: *Lactobacillus reuteri* DSM 17938 vs. placebo in children with acute diarrhoea - a double-blind study. *Alimentary Pharmacology & Therapeutics* **36**, 363-369.

Fukiya, S., Sugiyama, T., Kano, Y. & Yokota, A. (2010). Characterisation of an insertion sequence-like element, ISBlo15, identified in a size-increased cryptic plasmid pBK283 in *Bifidobacterium longum* BK28. *Journal of Bioscience and Bioengineering* **110**, 141-146.

Fukuda, S., Toh, H., Hase, K. Fukuda, S., Toh, H., Hase, K., Oshima, K., Nakanishi, Y., Yoshimura, K.,Tobe, T.,Clarke,J.M., Topping, D.L.,Suzuki, T., Taylor, T.D., Itoh, K., Kikuchi, J., Morita, H.,Hattori, M. & Ohno, H. (2011). Bifidobacteria can protect from enteropathogenic infection through production of acetate. *Nature* **469**, 543-547.

Garrigues, C., Johansen, E. & Pedersen, M. B. (2010). Complete Genome Sequence of *Bifidobacterium animalis* subsp. *lactis* BB-12, a Widely Consumed Probiotic Strain. *Journal of Bacteriology* **192**, 2467-2468.

Garro, M. S., de Giori, G. S., de Valdez, G. F. & Oliver, G. (1994). α-D-Galactosidase (EC 3.2.1.22) from *Bifidobacterium longum*. *Letters in Applied Microbiology* **19**, 16-19.

Garwin, J. L., Klages, A. L. & Cronan, J. E., Jr. (1980a). Beta-ketoacyl-acyl carrier protein synthase II of *Escherichia coli*. Evidence for function in the thermal regulation of fatty acid synthesis. *The Journal of Biological Chemistry* **255**, 3263-3265.

Garwin, J. L., Klages, A. L. & Cronan, J. E., Jr. (1980b). Structural, enzymatic, and genetic studies of beta-ketoacyl-acyl carrier protein synthases I and II of *Escherichia coli*. *The Journal of Biological Chemistry* **255**, 11949-11956.

Gibbs, M. J., Smeianov, V. V., Steele, J. L., Upcroft, P. & Efimov, B. A. (2006). Two families of rep-like genes that probably originated by interspecies recombination are represented in viral, plasmid, bacterial, and parasitic protozoan genomes. *Molecular Biology and Evolution* **23**, 1097-1100.

Gibson, G. R. & Roberfroid, M. B. (1995). Dietary modulation of the human colonic microbiota: introducing the concept of prebiotics. *The Journal of Nutrition* **125**, 1401-1412.

Gill, I. & Valivety, R. (1997). Polyunsaturated fatty acids, Part 1: Occurrence, biological activities and applications. *Trends in Biotechnology* **15**, 401-409.

Goulas, T., Goulas, A., Tzortzis, G. & Gibson, G. R. (2009). A novel alpha-galactosidase from *Bifidobacterium bifidum* with transgalactosylating properties: gene molecular cloning and heterologous expression. *Applied Microbiology and Biotechnology* **82**, 471-477.

Grabitske, H. A. & Slavin, J. L. (2009). Gastrointestinal effects of low-digestible carbohydrates. *Critical Reviews in Food Science and Nutrition* **49**, 327-360.

Griinari, J. M., Corl, B. A., Lacy, S. H., Chouinard, P. Y., Nurmela, K. V. & Bauman, D. E. (2000). Conjugated linoleic acid is synthesised endogenously in lactating dairy cows by Delta(9)-desaturase. *The Journal of Nutrition* **130**, 2285-2291.

Guarner, F. & Malagelada, J. R. (2003). Gut flora in health and disease. *Lancet* **361**, 512-519.

Gueimonde, M., Laitinen, K., Salminen, S. & Isolauri, E. (2007). Breast milk: a source of bifidobacteria for infant gut development and maturation? *Neonatology* **92**, 64-66.

Guglielmetti, S., Karp, M., Mora, D., Tamagnini, I. & Parini, C. (2007). Molecular characterisation of *Bifidobacterium longum* biovar *longum* NAL8 plasmids and construction of a novel replicon screening system. *Applied Microbiology and Biotechnology* **74**, 1053-1061.

Ha, Y. L., Grimm, N. K. & Pariza, M. W. (1987). Anticarcinogens from fried ground beef: heat-altered derivatives of linoleic acid. *Carcinogenesis* **8**, 1881-1887.

Ham, J.-S., Lee, T., Byun, M.J., Lee, K.T., Kim, M.K., Han, G.S., Jeong, S.G., Oh, M.H., Kim, D.H. & Kim, H. (2011). Complete Genome Sequence of *Bifidobacterium longum* subsp. *longum* KACC 91563. *Journal of Bacteriology* **193**, 5044.

Han, L., Lobo, S. & Reynolds, K. A. (1998). Characterisation of β-Ketoacyl-Acyl Carrier Protein Synthase III from *Streptomyces glaucescens* and Its Role in Initiation of Fatty Acid Biosynthesis. *Journal of Bacteriology* **180**, 4481-4486.

Hao, Y., Huang, D., Guo, H. Xiao, M., An, H., Zhao, L., Zuo, F., Zhang, B., Hu, S., Song, S., Chen, S. & Ren, F. (2011). Complete genome sequence of *Bifidobacterium longum* subsp. *longum* BBMN68, a new strain from a healthy chinese centenarian. *Journal of Bacteriology* **193**, 787-788.

Harfoot, C. G. & Hazlewood, G. P. (1997). Lipid metabolism in the rumen. In *The Rumen Microbial Ecosystem*, pp. 382-426. Edited by P. N. Hobson & C. S. Stewart: Springer Netherlands.

Hattori, M. & Taylor, T. D. (2009). The human intestinal microbiome: a new frontier of human biology. *DNA research : an international journal for rapid publication of reports on genes and genomes* **16**, 1-12.

Henrissat, B. (1991). A classification of glycosyl hydrolases based on amino acid sequence similarities. *The Biochemical Journal* **280**, 309-316.

Hinz, S. W., van den Brock, L. A., Beldman, G., Vincken, J. P. & Voragen, A. G. (2004). Beta-galactosidase from *Bifidobacterium adolescentis* DSM20083 prefers beta(1,4)-galactosides over lactose. *Applied Microbiology and Biotechnology* **66**, 276-284.

Hirayama, Y., Sakanaka, M., Fukuma, H., Murayama, H., Kano, Y., Fukiya, S. & Yokota, A. (2012). Development of a double-crossover markerless gene deletion system in *Bifidobacterium longum*: functional analysis of the alpha-galactosidase gene for raffinose assimilation. *Applied and Environmental Microbiology* **78**, 4984-4994.

Hold, G. L., Pryde, S. E., Russell, V. J., Furrie, E. & Flint, H. J. (2002). Assessment of microbial diversity in human colonic samples by 16S rDNA sequence analysis. *FEMS Microbiology Ecology* **39**, 33-39.

Hornung, E., Krueger, C., Pernstich, C., Gipmans, M., Porzel, A. & Feussner, I. (2005). Production of (10E,12Z)-conjugated linoleic acid in yeast and tobacco seeds. *Biochimica et biophysica acta* **1738**, 105-114.

Hsu, C. A., Lee, S. L. & Chou, C. C. (2007). Enzymatic production of galactooligosaccharides by beta-galactosidase from *Bifidobacterium longum* BCRC 15708. *Journal of Agricultural and Food Chemistry* **55**, 2225-2230.

Hung, M. N. & Lee, B. H. (2002). Purification and characterisation of a recombinant beta-galactosidase with transgalactosylation activity from *Bifidobacterium infantis* HL96. *Applied Microbiology and Biotechnology* **58**, 439-445.

Ishizeki, S., Sugita, M., Takata, M. & Yaeshima, T. (2013). Effect of administration of bifidobacteria on intestinal microbiota in low-birth-weight infants and transition of administered bifidobacteria: A comparison between one-species and three-species administration. *Anaerobe* **23**, 38-44.

Jacob, F. & Monod, J. (1961). Genetic regulatory mechanisms in the synthesis of proteins. *Journal of Molecular Biology* **3**, 318-356.

Jankovic, I., Sybesma, W., Phothirath, P., Ananta, E. & Mercenier, A. (2010). Application of probiotics in food products ,challenges and new approaches. *Current Opinion in Biotechnology* **21**, 175-181.

Jenkins, T. C., Wallace, R. J., Moate, P. J. & Mosley, E. E. (2008). Board-invited review: Recent advances in biohydrogenation of unsaturated fatty acids within the rumen microbial ecosystem. *Journal of Animal Science* **86**, 397-412.

Jiang, J., Bjorck, L. & Fonden, R. (1998). Production of conjugated linoleic acid by dairy starter cultures. *Journal of Applied Microbiology* **85**, 95-102.

Jie, Z., Bang-Yao, L., Ming-Jie, X., Hai-Wei, L., Zu-Kang, Z., Ting-Song, W. & Craig, S. A. (2000). Studies on the effects of polydextrose intake on physiologic functions in Chinese people. *The American Journal of Clinical Nutrition* **72**, 1503-1509.

Joo, Y. C., Seo, E. S., Kim, Y. S., Kim, K. R., Park, J. B. & Oh, D. K. (2012). Production of 10-hydroxystearic acid from oleic acid by whole cells of recombinant *Escherichia coli* containing oleate hydratase from *Stenotrophomonas maltophilia*. *Journal of Biotechnology* **158**, 17-23.

Jorgensen, F., Hansen, O. C. & Stougaard, P. (2001). High-efficiency synthesis of oligosaccharides with a truncated beta-galactosidase from *Bifidobacterium bifidum*. *Applied Microbiology and Biotechnology* **57**, 647-652.

Kass, L. R., Brock, D. J. H. & Bloch, K. (1967). β-Hydroxydecanoyl Thioester Dehydrase: I. purification and properties. *Journal of Biological Chemistry* **242**, 4418-4431.

Kepler, C. R., Hirons, K. P., McNeill, J. J. & Tove, S. B. (1966). Intermediates and Products of the Biohydrogenation of Linoleic Acid by *Butyrivibrio fibrisolvens*. *Journal of Biological Chemistry* **241**, 1350-1354.

Kepler, C. R., Tucker, W. P. & Tove, S. B. (1970). Biohydrogenation of Unsaturated Fatty Acids: iv. substrate specificity and inhibition of linoleate δ12-cis, δ11-trans-isomerase from *butyrivibrio fibrisolvens*. *Journal of Biological Chemistry* **245**, 3612-3620.

Kil, K. S., Cunningham, M. W. & Barnett, L. A. (1994). Cloning and sequence analysis of a gene encoding a 67-kilodalton myosin-cross-reactive antigen of *Streptococcus pyogenes* reveals its similarity with class II major histocompatibility antigens. *Infection and Immunity* **62**, 2440-2449.

Kim, J. F., Jeong, H., Yu, D. S., Choi, S.H., Hur, C.G., Park, M.S., Yoon, S.H., Kim, D.W., Ji, G.E., Park, H.S. & Oh, T.K. (2009). Genome Sequence of the Probiotic *Bacterium Bifidobacterium animalis* subsp. *lactis* AD011. *Journal of Bacteriology* **191**, 678-679.

Kingsbury, D. T. & Helinski, D. R. (1970). DNA polymerase as a requirement for the maintenance of the bacterial plasmid colicinogenic factor E1. *Biochemical and Biophysical Research Communications* **41**, 1538-1544.

Klaenhammer, T.R., Kleerebezem, M., Volkmar Kopp, V. & Rescigno, M. (2012). The impact of probiotics and prebiotics on the immune system. *Nature Reviews Immunology* **12,** 728-734.

Klijn, A., Moine, D., Delley, M., Mercenier, A., Arigoni, F. & Pridmore, R. D. (2006). Construction of a Reporter Vector for the Analysis of *Bifidobacterium longum* Promoters. *Applied and Environmental Microbiology* **72**, 7401-7405.

Koenig, J. E., Spor, A., Scalfone, N., Fricker, A. D., Stombaugh, J., Knight, R., Angenent, L. T. & Ley, R. E. (2011). Succession of microbial consortia in the developing infant gut microbiome. *Proceedings of the National Academy of Sciences of the United States of America* **108**, 4578-4585.

Lattimer, J. M. & Haub, M. D. (2010). Effects of Dietary Fiber and Its Components on Metabolic Health. *Nutrients* **2**, 1266-1289.

Leder, S., Hartmeier, W. & Marx, S. P. (1999). Alpha-galactosidase of *Bifidobacterium adolescentis* DSM 20083. *Current Oppinion in Microbiology* **38**, 101-106.

Lee, J.H., Karamychev, V., Kozyavkin, S., Mills, D., Pavlov, A.R., Pavlova, N.V., Polouchine, N.N., Richardson, P.M., Shakhova, V.V., Slesarev, A.I., Weimer, B. & O' Sullivan, D.J. (2008). Comparative genomic analysis of the gut bacterium *Bifidobacterium longum* reveals loci susceptible to deletion during pure culture growth. *BMC Genomics* **9**, 247.

Lee, J. H. & O'Sullivan, D. J. (2006). Sequence analysis of two cryptic plasmids from *Bifidobacterium longum* DJO10A and construction of a shuttle cloning vector. *Applied and Environmental Microbiology* **72**, 527-535.

Leon, R. K. & Bloch, K. (1967). On the Enzymatic Synthesis of Unsaturated Fatty Acids in *Escherichia coli*. *Proceedings of the National Academy of Sciences of the United States of America* **58**, 1168-1173.

Li, F., Zhang, X. & Lim, S. (2014). Molecular Cloning and Expression of the β-Galactosidase Gene from *Bifidobacterium longum* LL04 in *Escherichia coli*. In *Proceedings of the 2012 International Conference on Applied Biotechnology (ICAB 2012)*, pp. 1279-1290. Edited by T.-C. Zhang, P. Ouyang, S. Kaplan & B. Skarnes: Springer Berlin Heidelberg.

Li, Y., Florova, G. & Reynolds, K. A. (2005). Alteration of the Fatty Acid Profile of *Streptomyces coelicolor* by Replacement of the Initiation Enzyme 3-Ketoacyl Acyl Carrier Protein Synthase III (FabH). *Journal of Bacteriology* **187**, 3795-3799.

Lin, T. Y., Lin, C. W. & Wang, Y. J. (2002). Linoleic acid Isomerase Activity in Enzyme Extracts from *Lactobacillus Acidophilus* and *Propionibacterium Freudenreichii* ssp. *Shermanii*. *Journal of Food Science* **67**, 1502-1505.

Loquasto, J. R., Barrangou, R., Dudley, E. G., Stahl, B., Chen, C. & Roberts, R. F. (2013). *Bifidobacterium animalis* subsp. *lactis* ATCC 27673 Is a Genomically Unique Strain within Its Conserved Subspecies. *Applied and Environmental Microbiology* **79**, 6903-6910.

Lorca, G. L., Barabote, R. D., Zlotopolski, V., Tran, C., Winnen, B., Hvorup, R.N., Stonestrom, A.J., Nguyen, E., Huang, L.W., Kim, D.S. & Saier, M.H. Jr. (2007). Transport capabilities of eleven gram-positive bacteria: comparative genomic analyses. *Biochim Biophys Acta* **1768**, 1342-1366.

Luo, X., Zhang, L., Li, H., Zhang, S., Jiao, Y., Wang, S., Xue, C. & Fan, R. (2013). Comparison of enzymatic activity of two linoleic acid isomerases expressed in *E. coli*. *Molecular Biology Reports* **40**, 5913-5919.

Macfarlane, G. T., Steed, H. & Macfarlane, S. (2008). Bacterial metabolism and health-related effects of galacto-oligosaccharides and other prebiotics. *Journal of Applied Microbiology* **104**, 305-344.

Marco, M. L., Pavan, S. & Kleerebezem, M. (2006). Towards understanding molecular modes of probiotic action. *Current Opinion in Biotechnology* **17**, 204-210.

Martínez-Villaluenga, C. & Gómez, R. (2007). Characterisation of bifidobacteria as starters in fermented milk containing raffinose family of oligosaccharides from lupin as prebiotic. *International Dairy Journal* **17**, 116-122.

Matsumura, H., Takeuchi, A. & Kano, Y. (1997). Construction of *Escherichia coli-Bifidobacterium longum* shuttle vector transforming *B. longum* 105-A and 108-A. *Bioscience, Biotechnology, and Biochemistry* **61**, 1211-1212.

Maze, A., O'Connell-Motherway, M., Fitzgerald, G. F., Deutscher, J. & van Sinderen, D. (2007). Identification and characterisation of a fructose phosphotransferase system in *Bifidobacterium breve* UCC2003. *Applied and Environmental Microbiology* **73**, 545-553.

Milani, C., Duranti, S., Lugli, G. A., Bottacini, F., Strati, F., Arioli, S., Foroni, E., Turroni, F., van Sinderen, D. & Ventura, M. (2013). Comparative genomics of *Bifidobacterium animalis* subsp. *lactis* reveals a strict monophyletic bifidobacterial taxon. *Applied and Environmental Microbiology* **79**, 4304-4315.

Missich, R., Sgorbati, B. & LeBlanc, D. J. (1994). Transformation of *Bifidobacterium longum* with pRM2, a constructed *Escherichia coli-B. longum* shuttle vector. *Plasmid* **32**, 208-211.

Mohan, S., Kelly, T. M., Eveland, S. S., Raetz, C. R. & Anderson, M. S. (1994). An *Escherichia coli* gene (FabZ) encoding (3R)-hydroxymyristoyl acyl carrier protein dehydrase. Relation to fabA and suppression of mutations in lipid A biosynthesis. *The Journal of Biological Chemistry* **269**, 32896-32903.

Moller, P. L., Jorgensen, F., Hansen, O. C., Madsen, S. M. & Stougaard, P. (2001). Intra- and extracellular beta-galactosidases from *Bifidobacterium bifidum* and *B. infantis*: molecular cloning, heterologous expression, and comparative characterisation. *Applied and Environmental Microbiology* **67**, 2276-2283.

Moon, G. S., Wegmann, U., Gunning, A. P., Gasson, M. J. & Narbad, A. (2009). Isolation and characterisation of a theta-type cryptic plasmid from *Bifidobacterium longum* FI10564. *Journal of Microbiology and Biotechnology* **19**, 403-408.

Mulugu, S., Potnis, A., Shamsuzzaman, Taylor, J., Alexander, K. & Bastia, D. (2001). Mechanism of termination of DNA replication of *Escherichia coli* involves helicase–contrahelicase interaction. *Proceedings of the National Academy of Sciences* **98**, 9569-9574.

O'Connell, K. J., Motherway, M. O., Hennessey, A. A., Brodhun, F., Ross, R. P., Feussner, I., Stanton, C., Fitzgerald, G. F. & van Sinderen, D. (2013a). Identification and characterisation of an oleate hydratase-encoding gene from *Bifidobacterium breve*. *Bioengineered* **4**, 313-321.

O'Connell, K. J., O'Connell Motherway, M., O'Callaghan, J., Fitzgerald, G. F., Ross, R. P., Ventura, M., Stanton, C. & van Sinderen, D. (2013b). Metabolism of four α-glycosidic linkage-containing oligosaccharides by *Bifidobacterium breve* UCC2003. *Applied and Environmental Microbiology*. **79**, 6280-6292.

O'Connell Motherway, M., Fitzgerald, G. F., Neirynck, S., Ryan, S., Steidler, L. & van Sinderen, D. (2008). Characterisation of ApuB, an extracellular type II amylopullulanase from *Bifidobacterium breve* UCC2003. *Applied and Environmental Microbiology* **74**, 6271-6279.

O'Connell Motherway, M., Fitzgerald, G. F. & van Sinderen, D. (2011a). Metabolism of a plant derived galactose-containing polysaccharide by *Bifidobacterium breve* UCC2003. *Microbial Biotechnology* **4**, 403-416.

O'Connell Motherway, M., Zomer, A., Leahy, S. C., Reunanen, J., Bottacini, F., Claesson, M.J., O'Brien, F., Flynn, K., Casey, P.G., Munoz, J.A., Kearney, B., Houston, A.M., O' Mahony, C., Higgins, D.G., Shanahan, F., Palva, A., de Vos, W.M., Fitzgerald, G.F., Ventura, M., O'Toole, P.W. & van Sinderen, D. (2011b). Functional genome analysis of *Bifidobacterium breve* UCC2003 reveals

type IVb tight adherence (Tad) pili as an essential and conserved host-colonisation factor. *Proceedings of the National Academy of Sciences of the United States of America* **108**, 11217-11222.

O'Connell Motherway, M., Kinsella, M., Fitzgerald, G. F. & van Sinderen, D. (2013). Transcriptional and functional characterisation of genetic elements involved in galacto-oligosaccharide utilisation by *Bifidobacterium breve* UCC2003. *Microbial Biotechnology* **6**, 67-79.

O'Hara, A. M. & Shanahan, F. (2007). Mechanisms of action of probiotics in intestinal diseases. *The Scientific World Journal* **7**, 31-46.

O'Riordan, K. & Fitzgerald, G. F. (1999). Molecular characterisation of a 5.75-kb cryptic plasmid from *Bifidobacterium breve* NCFB 2258 and determination of mode of replication. *FEMS Microbiology Letters* **174**, 285-294.

Ogawa, J., Matsumura, K., Kishino, S., Omura, Y. & Shimizu, S. (2001). Conjugated linoleic acid accumulation via 10-hydroxy-12-octadecaenoic acid during microaerobic transformation of linoleic acid by *Lactobacillus acidophilus*. *Applied and Environmental Microbiology* **67**, 1246-1252.

Ogawa, J., Kishino, S., Ando, A., Sugimoto, S., Mihara, K. & Shimizu, S. (2005). Production of conjugated fatty acids by lactic acid bacteria. *Journal of Bioscience and Bioengineering* **100**, 355-364.

Ouwehand, A. C. (2007). Antiallergic Effects of Probiotics. *The Journal of Nutrition* **137**, 794S-797S.

Palframan, R. J., Gibson, G. R. & Rastall, R. A. (2003). Carbohydrate preferences of *Bifidobacterium* species isolated from the human gut. *Current Issues in Intestinal Microbiology* **4**, 71-75.

Parche, S., Beleut, M., Rezzonico, E., Jacobs, D., Arigoni, F., Titgemeyer, F. & Jankovic, I. (2006). Lactose-over-glucose preference in *Bifidobacterium longum* NCC2705: *glcP*, encoding a glucose transporter, is subject to lactose repression. *Journal of Bacteriology* **188**, 1260-1265.

Parche, S., Amon, J., Jankovic, I., Rezzonico, E., Beleut, M., Barutçu, H., Schendel, I., Eddy, M.P., Burkovski, A., Arigoni, F. & Titgemeyer, F. (2007). Sugar transport systems of *Bifidobacterium longum* NCC2705. *Journal of Molecular Microbiology and Biotechnology* **12**, 9-19.

Park, M. S., Lee, K. H. & Ji, G. E. (1997a). Isolation and characterisation of two plasmids from *Bifidobacterium longum*. *Letters in Applied Microbiology* **25**, 5-7.

Park, M. S., Shin, D. W., Lee, K. H. & Ji, G. E. (1999). Sequence analysis of plasmid pKJ50 from *Bifidobacterium longum*. *Microbiology* **145**, 585-592.

Park, Y., Albright, K., Liu, W., Storkson, J., Cook, M. & Pariza, M. (1997b). Effect of conjugated linoleic acid on body composition in mice. *Lipids* **32**, 853-858.

Park, Y. S., Kim, K. H., Park, J. H., Oh, I. K. & Yoon, S. S. (2008). Isolation and molecular characterisation of a cryptic plasmid from *Bifidobacterium longum*. *Biotechnology Letters* **30**, 145-151.

Plumbridge, J. (2001). Regulation of PTS gene expression by the homologous transcriptional regulators, *Mlc* and *NagC*, in *Escherichia coli* (or how two similar repressors can behave differently). *Journal of Molecular Microbiology and Biotechnology* **3**, 371-380.

Plumbridge, J. A., Cochet, O., Souza, J. M., Altamirano, M. M., Calcagno, M. L. & Badet, B. (1993). Coordinated regulation of amino sugar-synthesising and -degrading enzymes in *Escherichia coli* K-12. *Journal of Bacteriology* **175**, 4951-4956.

Pokusaeva, K., O'Connell-Motherway, M., Zomer, A., Fitzgerald, G. F. & van Sinderen, D. (2009). Characterisation of two novel alpha-glucosidases from

Bifidobacterium breve UCC2003. *Applied and Environmental Microbiology* **75**, 1135-1143.

Pokusaeva, K., Neves, A. R., Zomer, A., O'Connell-Motherway, M., MacSharry, J., Curley, P., Fitzgerald, G. F. & van Sinderen, D. (2010). Ribose utilisation by the human commensal *Bifidobacterium breve* UCC2003. *Microbial Biotechnology* **3**, 311-323.

Pokusaeva, K., Fitzgerald, G. & Sinderen, D. (2011a). Carbohydrate metabolism in Bifidobacteria. *Genes and Nutrition* **6**, 285-306.

Pokusaeva, K., O'Connell-Motherway, M., Zomer, A., Macsharry, J., Fitzgerald, G. F. & van Sinderen, D. (2011b). Cellodextrin utilisation by *Bifidobacterium breve* UCC2003. *Applied and Environmental Microbiology* **77**, 1681-1690.

Polakis, S. E., Guchhait, R. B., Zwergel, E. E., Lane, M. D. & Cooper, T. G. (1974). Acetyl coenzyme A carboxylase system of *Escherichia coli*. Studies on the mechanisms of the biotin carboxylase- and carboxyltransferase-catalysed reactions. *The Journal of Biological Chemistry* **249**, 6657-6667.

Polan, C. E., McNeill, J. J. & Tove, S. B. (1964). Biohydrogenation of unsaturated fatty acids by rumen bacteria. *Journal of Bacteriology* **88**, 1056-1064.

Postma, P. W., Lengeler, J. W. & Jacobson, G. R. (1993). Phosphoenolpyruvate:carbohydrate phosphotransferase systems of bacteria. *Microbiological Reviews* **57**, 543-594.

Prescott, D. M. & Kuempel, P. L. (1972). Bidirectional Replication of the Chromosome in *Escherichia coli*. *Proceedings of the National Academy of Sciences* **69**, 2842-2845.

Punta, M., Coggill, P. C., Eberhardt, R. Y., Mistry, J., Tate, J., Boursnell, C., Pang, N., Forslund, K., Ceric, G., Clements, J., Heger, A., Holm, L., Sonnhammer, E.L., Eddy, S.R., Bateman, A. & Finn, RD. (2012). The Pfam protein families database. *Nucleic Acids Research* **40**, D290-301.

Quigley, E. M. M. (2010). Prebiotics and probiotics; modifying and mining the microbiota. *Pharmacological Research* **61**, 213-218.

Rabiu, B. A., Jay, A. J., Gibson, G. R. & Rastall, R. A. (2001). Synthesis and fermentation properties of novel galacto-oligosaccharides by beta-galactosidases from *Bifidobacterium* species. *Applied and Environmental Microbiology* **67**, 2526-2530.

Revill, W. P., Bibb, M. J. & Hopwood, D. A. (1995). Purification of a malonyltransferase from *Streptomyces coelicolor* A3(2) and analysis of its genetic determinant. *Journal of Bacteriology* **177**, 3946-3952.

Revill, W. P., Bibb, M. J. & Hopwood, D. A. (1996). Relationships between fatty acid and polyketide synthases from *Streptomyces coelicolor* A3(2): characterisation of the fatty acid synthase acyl carrier protein. *Journal of Bacteriology* **178**, 5660-5667.

Rock, C. O. & Cronan, J. E. (1996). *Escherichia coli* as a model for the regulation of dissociable (type II) fatty acid biosynthesis. *Biochimica et biophysica acta* **1302**, 1-16.

Rosberg-Cody, E., Johnson, M. C., Fitzgerald, G. F., Ross, P. R. & Stanton, C. (2007). Heterologous expression of linoleic acid isomerase from *Propionibacterium acnes* and anti-proliferative activity of recombinant trans-10, cis-12 conjugated linoleic acid. *Microbiology* **153**, 2483-2490.

Rosberg-Cody, E., Liavonchanka, A., Gobel, C., Ross, R. P., O'Sullivan, O., Fitzgerald, G. F., Feussner, I. & Stanton, C. (2011a). Myosin-cross-reactive antigen (MCRA) protein from *Bifidobacterium breve* is a FAD-dependent fatty acid hydratase which has a function in stress protection. *BMC Biochemistry* **12**, 9.

Rosberg-Cody, E., Stanton, C., O'Mahony, L., Wall, R., Shanahan, F., Quigley, E. M., Fitzgerald, G. F. & Ross, R. P. (2011b). Recombinant *lactobacilli* expressing linoleic acid isomerase can modulate the fatty acid composition of host adipose tissue in mice. *Microbiology* **157**, 609-615.

Rossi, A. & Corcoran, J. W. (1973). Identification of a multienzyme complex synthesising fatty acids in the actinomycete *Streptomyces erythreus*. *Biochemical and Biophysical Research Communications* **50**, 597-602.

Rossi, M., Brigidi, P., Gonzalez Vara y Rodriguez, A. & Matteuzzi, D. (1996). Characterisation of the plasmid pMB1 from *Bifidobacterium longum* and its use for shuttle vector construction. *Research in Microbiology* **147**, 133-143.

Rossi, M., Brigidi, P. & Matteuzzi, D. (1998). Improved cloning vectors for *Bifidobacterium* spp. *Letters in Applied Microbiology* **26**, 101-104.

Rossi, M., Altomare, L., Gonzalez Vara y Rodriguez, A., Brigidi, P. & Matteuzzi, D. (2000). Nucleotide sequence, expression and transcriptional analysis of the *Bifidobacterium longum* MB 219 *lacZ* gene. *Archives of microbiology* **174**, 74-80.

Rosson, R. A., Grund, A., Deng, M. & Sanchez-Riera, F. (2001). Linoleate isomerase. *World Patent* **100846**, 30.

Ruch, F. E. & Vagelos, P. R. (1973). The isolation and general properties of *Escherichia coli* malonyl coenzyme A-acyl carrier protein transacylase. *The Journal of Biological Chemistry* **248**, 8086-8094.

Ruiz-Moyano, S., Totten, S. M., Garrido, D. A., Smilowitz, J. T., German, J. B., Lebrilla, C. B. & Mills, D. A. (2013). Variation in Consumption of Human Milk Oligosaccharides by Infant Gut-Associated Strains of *Bifidobacterium breve*. *Applied and Environmental Microbiology* **79**, 6040-6049.

Ryan, S. M., Fitzgerald, G. F. & van Sinderen, D. (2005). Transcriptional Regulation and Characterisation of a Novel β-Fructofuranosidase-Encoding Gene from *Bifidobacterium breve* UCC2003. *Applied and Environmental Microbiology* **71**, 3475-3482.

Saier, M. H., Jr. & Ramseier, T. M. (1996). The catabolite repressor/activator (Cra) protein of enteric bacteria. *Journal of Bacteriology* **178**, 3411-3417.

Salminen, S., Bouley, C., Boutron-Ruault, M. C., Cummings, J..H, Franck, A., Gibson, G.R., Isolauri, E., Moreau, M.C., Roberfroid, M. & Rowland, I. (1998). Functional food science and gastrointestinal physiology and function. *The British Journal of Nutrition* **80**, S147-171.

Salminen, S. J., Gueimonde, M. & Isolauri, E. (2005). Probiotics that modify disease risk. *The Journal of Nutrition* **135**, 1294-1298.

Sangrador-Vegas, A., Stanton, C., van Sinderen, D., Fitzgerald, G. F. & Ross, R. P. (2007). Characterisation of plasmid pASV479 from *Bifidobacterium pseudolongum* subsp. *globosum* and its use for expression vector construction. *Plasmid* **58**, 140-147.

Savage, D. C. (2001). Microbial biota of the human intestine: a tribute to some pioneering scientists. *Current Issues in Intestinal Microbiology* **2**, 1-15.

Scardovi, V. & Trovatelli, I. D. (1965). The fructose-6-phosphate shunt as a peculiar pattern of hexose degradation in the genus *Bifidobacterium*. *Annals of Microbiology* **15**, 19-29.

Scheler, A. & Hillen, W. (1994). Regulation of xylose utilisation in *Bacillus licheniformis*: *Xyl* repressor-xyl-operator interaction studied by DNA modification protection and interference. *Molecular Microbiology* **13**, 505-512.

Schell, M. A., Karmirantzou, M., Snel, B., Vilanova, D., Berger, B., Pessi, G., Zwahlen, M.C., Desiere, F., Bork, P., Delley, M., Pridmore, R.D. & Arigoni, F. (2002). The genome sequence of *Bifidobacterium longum* reflects its adaptation to

the human gastrointestinal tract. *Proceedings of the National Academy of Sciences of the United States of America* **99**, 14422-14427.

Schiefner, A., Gerber, K., Seitz, S., Welte, W., Diederichs, K. & Boos, W. (2005). The Crystal Structure of *Mlc*, a Global Regulator of Sugar Metabolism in *Escherichia coli*. *Journal of Biological Chemistry* **280**, 29073-29079.

Schweizer, E. & Hofmann, J. (2004). Microbial type I fatty acid synthases (FAS): major players in a network of cellular FAS systems. *Microbiology and Molecular Biology Reviews* **68**, 501-517.

Sela, D. A., Chapman, J., Adeuya, A., Kim, J.H., Chen, F., Whitehead, T.R., Lapidus, A., Rokhsar, D.S., Lebrilla, C.B., German, J.B., Price, N.P., Richardson, P.M. & Mills, D.A. (2008). The genome sequence of *Bifidobacterium longum* subsp. *infantis* reveals adaptations for milk utilisation within the infant microbiome. *Proceedings of the National Academy of Sciences of the United States of America* **105**, 18964-18969.

Sela, D. A., Garrido, D., Lerno, L., Wu, S., Tan, K., Eom, H.-J., Joachimiak, A., Lebrilla, C. B. & Mills, D. A. (2012). *Bifidobacterium longum* subsp. *infantis* ATCC 15697 α-Fucosidases Are Active on Fucosylated Human Milk Oligosaccharides. *Applied and Environmental Microbiology* **78**, 795-803.

Sgorbati, B., Scardovi, V. & Leblanc, D. J. (1982). Plasmids in the Genus *Bifidobacterium*. *Journal of General Microbiology* **128**, 2121-2131.

Shkoporov, A. N., Efimov, B. A., Khokhlova, E. V., Steele, J. L., Kafarskaia, L. I. & Smeianov, V. V. (2008). Characterisation of plasmids from human infant *Bifidobacterium* strains: sequence analysis and construction of *E. coli*-*Bifidobacterium* shuttle vectors. *Plasmid* **60**, 136-148.

Silvester, K. R., Englyst, H. N. & Cummings, J. H. (1995). Ileal recovery of starch from whole diets containing resistant starch measured *in vitro* and fermentation of ileal effluent. *The American Journal of Clinical Nutrition* **62**, 403-411.

Simpson, P. J., Stanton, C., Fitzgerald, G. F. & Ross, R. P. (2003). Genomic diversity and relatedness of bifidobacteria isolated from a porcine cecum. *Journal of Bacteriology* **185**, 2571-2581.

Souza, T. C., Silva, A. M., Drews, J. R. P., Gomes, D. A., Vinderola, C. G. & Nicoli, J. R. (2013). In vitro evaluation of *Bifidobacterium* strains of human origin for potential use in probiotic functional foods. *Beneficial Microbes* **4**, 179-186.

Stahl, B. & Barrangou, R. (2012). Complete genome sequences of probiotic strains *Bifidobacterium animalis* subsp. *lactis* B420 and Bi-07. *Journal of Bacteriology* **194**, 4131-4132.

Stephen, A. M., Wiggins, H. S. & Cummings, J. H. (1987). Effect of changing transit time on colonic microbial metabolism in man. *Gut* **28**, 601-609.

Summers, R. G., Ali, A., Shen, B., Wessel, W. A. & Hutchinson, C. R. (1995). Malonyl-Coenzyme A:Acyl Carrier Protein Acyltransferase of *Streptomyces glaucescens*: A Possible Link between Fatty Acid and Polyketide Biosynthesis. *Biochemistry* **34**, 9389-9402.

Sun, Z., Chen, X., Wang, J. Gao, P., Zhou, Z., Ren, Y., Sun, T., Wang, L., Meng, H., Chen, W. & Zhang H (2010). Complete genome sequence of probiotic *Bifidobacterium animalis* subsp. *lactis* strain V9. *Journal of Bacteriology* **192**, 4080-4081.

Takahata, M., Toh, H., Nakano, A., Takagi, M., Murakami, M., Ishii, Y., Takizawa, T., Tanabe, S. & Morita, H. (2013). Complete sequence analysis of two cryptic plasmids from *Bifidobacterium kashiwanohense* JCM 15439 (type strain) isolated from healthy infant feces. *Animal Science Journal* **85**, 158-63.

Takechi, S. & Itoh, T. (1995). Initiation of unidirectional CoIE2 DNA replication by a unique priming mechanism. *Nucleic Acids Research* **23**, 4196-4201.

Takechi, S., Matsui, H. & Itoh, T. (1995). Primer RNA synthesis by plasmid-specified Rep protein for initiation of ColE2 DNA replication. *The EMBO Journal* **14**, 5141-5147.

Tanaka, K., Samura, K. & Kano, Y. (2005). Structural and functional analysis of pTB6 from *Bifidobacterium longum*. *Bioscience, Biotechnology, and Biochemistry* **69**, 422-425.

Tannock, G. W. (2001). Molecular assessment of intestinal microflora. *The American Journal of Clinical Nutrition* **73**, 410S-414S.

Tissier, H. (1900). Recherchers sur la flora intestinale normale pathologique du nourisson. *Thesis, University of Paris, Paris France.*

Titgemeyer, F., Reizer, J., Reizer, A. & Saier, M. H. (1994). Evolutionary relationships between sugar kinases and transcriptional repressors in bacteria. *Microbiology* **140**, 2349-2354.

Topping, D. L. & Clifton, P. M. (2001). Short-chain fatty acids and human colonic function: roles of resistant starch and nonstarch polysaccharides. *Physiological Reviews* **81**, 1031-1064.

Trojanova, I., Vlkova, E., Rada, V. & Marounek, M. (2006). Different utilisation of glucose and raffinose in *Bifidobacterium breve* and *Bifidobacterium animalis*. *Folia microbiologica* **51**, 320-324.

Tsay, J. T., Oh, W., Larson, T. J., Jackowski, S. & Rock, C. O. (1992). Isolation and characterisation of the beta-ketoacyl-acyl carrier protein synthase III gene (fabH) from *Escherichia coli* K-12. *The Journal of Biological Chemistry* **267**, 6807-6814.

Tuohy, K. M., Ziemer, C. J., Klinder, A., Knöbel, Y., Pool-Zobel, B. L. & Gibson, G. R. (2002). A Human Volunteer Study to Determine the Prebiotic Effects of Lactulose Powder on Human Colonic Microbiota. *Microbial Ecology in Health and Disease* **14**, 165-173.

Turroni, F., Ribbera, A., Foroni, E., van Sinderen, D. & Ventura, M. (2008). Human gut microbiota and bifidobacteria: from composition to functionality. *Antonie van Leeuwenhoek* **94**, 35-50.

Turroni, F., Bottacini, F., Foroni, E. Mulder, I., Kim, J.H., Zomer, A., Sánchez, B., Bidossi, A., Ferrarini, A., Giubellini, V., Delledonne, M., Henrissat, B., Coutinho, P., Oggioni, M., Fitzgerald, G.F., Mills, D., Margolles, A., Kelly, D., van Sinderen, D. & Ventura, M. (2010). Genome analysis of *Bifidobacterium bifidum* PRL2010 reveals metabolic pathways for host-derived glycan foraging. *Proceedings of the National Academy of Sciences of the United States of America* **107**, 19514-19519.

Turroni, F., Peano, C., Pass, D. A., Foroni, E., Severgnini, M., Claesson, M. J., Kerr, C., Hourihane, J., Murray, D., Fuligni, F., Gueimonde, M., Margolles, A., De Bellis, G., O'Toole, P. W., van Sinderen, D., Marchesi, J. R., & Ventura, M. (2012a). Diversity of bifidobacteria within the infant gut microbiota. *PLoS One* **7**, e36957.

Turroni, F., Peano , C., Pass ,D.A., Foroni, E., Severgnini, M., Claesson, M.J., Kerr, C., Hourihane, J., Murray, D., Fuligni, F., Gueimonde, M., Margolles, A., De Bellis, G., O'Toole, P.W. , van Sinderen, D., Marchesi, J.R. & Ventura ,M. (2012b). Diversity of Bifidobacteria within the Infant Gut Microbiota. *PLoS ONE* **7**, e36957.

Turroni, F., Strati, F., Foroni, E., Serafini, F., Duranti, S., van Sinderen, D. & Ventura, M. (2012c). Analysis of Predicted Carbohydrate Transport Systems Encoded by *Bifidobacterium bifidum* PRL2010. *Applied and Environmental Microbiology* **78**, 5002-5012.

Vaishampayan, P. A., Kuehl, J. V., Froula, J. L., Morgan, J. L., Ochman, H. & Francino, M. P. (2010). Comparative metagenomics and population dynamics of the gut microbiota in mother and infant. *Genome Biology and Evolution* **2**, 53-66.

van den Broek, L. A., Struijs, K., Verdoes, J. C., Beldman, G. & Voragen, A. G. (2003). Cloning and characterisation of two alpha-glucosidases from

Bifidobacterium adolescentis DSM20083. *Applied Microbiology and Biotechnology* **61**, 55-60.

Van Laere, K. M., Hartemink, R., Beldman, G., Pitson, S., Dijkema, C., Schols, H. A. & Voragen, A. G. (1999). Transglycosidase activity of *Bifidobacterium adolescentis* DSM 20083 alpha-galactosidase. *Applied Microbiology and Biotechnology* **52**, 681-688.

Ventura, M., Canchaya, C., Del Casale, A., Dellaglio, F., Neviani, E., Fitzgerald, G. F. & van Sinderen, D. (2006). Analysis of bifidobacterial evolution using a multilocus approach. *International Journal of Systematic and Evolutionary Microbiology* **56**, 2783-2792.

Ventura, M., Canchaya, C., Tauch, A., Chandra, G., Fitzgerald, G. F., Chater, K. F. & van Sinderen, D. (2007a). Genomics of Actinobacteria: Tracing the Evolutionary History of an Ancient Phylum. *Microbiology and Molecular Biology Reviews* **71**, 495-548.

Ventura, M., O'Connell-Motherway, M., Leahy, S., Moreno-Munoz, J. A., Fitzgerald, G. F. & van Sinderen, D. (2007b). From bacterial genome to functionality; case bifidobacteria. *International Journal of Food Microbiology* **120**, 2-12.

Ventura, M., Turroni, F., Zomer, A., Foroni, E., Giubellini, V., Bottacini, F., Canchaya, C., Claesson, M.J., He, F., Mantzourani, M., Mulas, L., Ferrarini, A., Gao, B., Delledonne, M., Henrissat, B., Coutinho, P., Oggioni, M., Gupta, R.S., Zhang, Z., Beighton, D., Fitzgerald, G.F., O' Toole, P.W. & van Sinderen, D. (2009). The *Bifidobacterium dentium* Bd1 genome sequence reflects its genetic adaptation to the human oral cavity. *PLoS genetics* **5**, e1000785.

Ventura, M., Turroni, F., Motherway, M. O., MacSharry, J. & van Sinderen, D. (2012). Host-microbe interactions that facilitate gut colonisation by commensal bifidobacteria. *Trends in Microbiology* **20**, 467-476.

Ventura, M., Turroni, F., Lugli, G. A. & van Sinderen, D. (2014). Bifidobacteria and humans: our special friends, from ecological to genomics perspectives. *Journal of the Science of Food and Agriculture* **94**, 163-168.

Viborg, A. H., Katayama, T., Hachem, M. A., Andersen, M. C. F., Nishimoto, M., Clausen, M. H., Urashima, T., Svensson, B. & Kitaoka, M. (2013). Distinct substrate specificities of three glycoside hydrolase family 42 β-galactosidases from *Bifidobacterium longum* subsp. *infantis* ATCC 15697. *Glycobiology* **00,** 1–9

Volkov, A., Liavonchanka, A., Kamneva, O., Fiedler, T., Goebel, C., Kreikemeyer, B. & Feussner, I. (2010). Myosin cross-reactive antigen of *Streptococcus pyogenes* M49 encodes a fatty acid double bond hydratase that plays a

role in oleic acid detoxification and bacterial virulence. *The Journal of Biological Chemistry* **285**, 10353-10361.

Wakinaka, T., Kiyohara, M., Kurihara, S., Hirata, A., Chaiwangsri, T., Ohnuma, T., Fukamizo, T., Katayama, T., Ashida, H. & Yamamoto, K. (2013). Bifidobacterial alpha-galactosidase with unique carbohydrate-binding module specifically acts on blood group B antigen. *Glycobiology* **23**, 232-240.

Wall, R., Ross, R. P., Fitzgerald, G. F. & Stanton, C. (2010). Fatty acids from fish: the anti-inflammatory potential of long-chain omega-3 fatty acids. *Nutrition Reviews* **68**, 280-289.

Walton, G. E., van den Heuvel, E. G., Kosters, M. H., Rastall, R. A., Tuohy, K. M. & Gibson, G. R. (2012). A randomised crossover study investigating the effects of galacto-oligosaccharides on the faecal microbiota in men and women over 50 years of age. *The British Journal of Nutrition* **107**, 1466-1475.

Weber, F. J., Isken, S. & de Bont, J. A. (1994). Cis/trans isomerisation of fatty acids as a defence mechanism of *Pseudomonas putida* strains to toxic concentrations of toluene. *Microbiology* **140**, 2013-2017.

Weeks, G. & Wakil, S. J. (1968). Studies on the Mechanism of Fatty Acid Synthesis: XVIII. preparation and general properties of the enoyl acyl carrier protein reductases from *Escherichia coli*. *Journal of Biological Chemistry* **243**, 1180-1189.

Wei, Y. X., Zhang, Z. Y., Liu, C., Zhu, Y. Z., Zhu, Y. Q., Zheng, H., Zhao, G. P., Wang, S. & Guo, X. K. (2010). Complete genome sequence of *Bifidobacterium longum* JDM301. *Journal of Bacteriology* **192**, 4076-4077.

White, S. W., Zheng, J., Zhang, Y. M. & Rock (2005). The structural biology of type II fatty acid biosynthesis. *Annual Review of Biochemistry* **74**, 791-831.

Wintjens, R. & Rooman, M. (1996). Structural Classification of HTH DNA-binding Domains and Protein-DNA Interaction Modes. *Journal of Molecular Biology* **262**, 294-313.

Yang, B., Chen, H., Song, Y., Chen, Y. Q., Zhang, H. & Chen, W. (2012). Myosin-cross-reactive antigens from four different lactic acid bacteria are fatty acid hydratases. *Biotechnology Letters*. **35,** 75-81.

Yasueda, H., Horii, T. & Itoh, T. (1989). Structural and functional organisation of ColE2 and ColE3 replicons. *Molecular & General genetics* **215**, 209-216.

Yi, S. H., Alli, I., Park, K. H. & Lee, B. (2011). Overexpression and characterisation of a novel transgalactosylic and hydrolytic β-galactosidase from a human isolate *Bifidobacterium breve* B24. *New Biotechnology* **28**, 806-813.

Yoshida, E., Sakurama, H., Kiyohara, M. Nakajima, M., Kitaoka, M., Ashida, H., Hirose, J., Katayama, T., Yamamoto, K. & Kumagai, H. (2012). *Bifidobacterium longum* subsp. *infantis* uses two different β-galactosidases for selectively degrading type-1 and type-2 human milk oligosaccharides. *Glycobiology* **22**, 361-368.

Zhao, H., Lu, L., Xiao, M., Wang, Q., Lu, Y., Liu, C., Wang, P., Kumagai, H. & Yamamoto, K. (2008). Cloning and characterisation of a novel α-galactosidase from *Bifidobacterium breve* 203 capable of synthesising Gal-α-1,4 linkage. *FEMS Microbiology Letters* **285**, 278-283.

Zheng, Y., Yin, H., Boeglin, W. E., Elias, P. M., Crumrine, D., Beier, D. R. & Brash, A. R. (2011). Lipoxygenases Mediate the Effect of Essential Fatty Acid in Skin Barrier Formation: a proposed role in releasing omega-hydroxyceramide for construction of the corneocyte lipid envelope. *Journal of Biological Chemistry* **286**, 24046-24056.

Zoetendal, E. G., Akkermans, A. D. & De Vos, W. M. (1998). Temperature gradient gel electrophoresis analysis of 16S rRNA from human fecal samples reveals

stable and host-specific communities of active bacteria. *Applied and Environmental Microbiology* **64**, 3854-3859.

Chapter II

Identification and characterisation of an oleate hydratase-encoding gene from *Bifidobacterium breve*.

Alan A. Hennessey performed GLC analysis.

Florian Brodhun and Ivo Feussner performed hydratase analysis.

This chapter was published in:

O'Connell, K. J., O'Connell Motherway, M., Hennessey, A. A., Brodhun, F., Ross, R. P., Feussner, I., Stanton, C., Fitzgerald, G. F. & van Sinderen, D. (2013). Identification and characterisation of an oleate hydratase-encoding gene from *Bifidobacterium breve*. *Bioengineered* **4,** 313–321.

2.1 ABSTRACT

Bifidobacteria are common commensals of the mammalian gastrointestinal tract. Previous studies have suggested that a bifidobacterial myosin cross reactive antigen (MCRA) protein plays a role in bacterial stress tolerance, while this protein has also been linked to the biosynthesis of conjugated linoleic acid (CLA) in bifidobacteria. In order to increase our understanding on the role of MCRA protein in bifidobacteria we created and analysed an insertion mutant of the MCRA-encoding gene of *B. breve* NCFB 2258. Our results demonstrate that the MCRA protein of *B. breve* NCFB 2258 does not appear to play a role in CLA production yet is an oleate hydratase, which contributes to bifidobacterial solvent stress protection.

2.2 INTRODUCTION

Bifidobacteria are Gram positive, non-motile, non-sporulating, anaerobic rods which possess a high G+C genome content, and belong to the phylum *Actinobacteria* and the family *Bifidobacteriaceae* (Ventura *et al.*, 2004; Ventura *et al.*, 2007). Bifidobacteria were first isolated more than a century ago (Tissier, 1900), and naturally inhabit the gastrointestinal tract of humans and other mammals, being particularly abundant in breast-fed infants (Turroni *et al.*, 2008; Turroni *et al.*, 2012). Bifidobacteria have been claimed to maintain gastrointestinal health (Tannock, 1997), and are therefore used as a health-promoting or probiotic bacterial ingredients in certain functional foods (Stanton *et al.*, 2005). Their reported beneficial effects on the host include inhibition of pathogens, alleviating lactose intolerance, enhancing natural immunity and reducing serum cholesterol (Leahy *et al.*, 2005; Liu *et al.*, 2007; Ventura *et al.*, 2012)

Bifidobacteria have been shown to convert oleic acid (OA) to 10-hydroxystearic acid (10-HSA) (Rosberg-Cody *et al.*, 2011a). OA is a mono-unsaturated fatty acid that occurs naturally in vegetable oils (Joo *et al.*, 2012), and is reported to confer a range of health benefits including the inhibition of adrenoleukodystrophy (ALD), a metabolic disorder that leads to demyelination in the central and peripheral nervous system (Perusi *et al.*, 1999), and reduction in blood pressure (Terés *et al.*, 2008). Microbial hydration of an unsaturated fatty acid was first reported for *Pseudomonas* sp. 3266 (now known as *Elizabethkingia meningoseptica*), which was shown to possess the ability to convert OA to 10-HSA (Wallen *et al.*, 1962). However, the isolation and biochemical characterisation of the enzyme responsible for this conversion, the oleate hydratase, was not reported until relatively recently (Bevers *et al.*, 2009). The latter authors also showed that expression of the *E. meningoseptica*

OA hydratase, was strongly up-regulated when this bacterium was grown in the presence of 0.3 % OA. Unsaturated fatty acids such as OA have been shown to be toxic to many bacteria (Raychowdhury *et al.*, 1985), which is probably due to the disruptive effect of these hydrophobic compounds on the bacterial membrane (Greenway & Dyke, 1979), combined with their inhibitory action on enoyl-ACP reductase (FabI) and thus fatty acid biosynthesis (Zheng *et al.*, 2005). The hydration of unsaturated fatty acids is therefore hypothesised to represent a detoxification mechanism in bacteria to aid survival in environments rich in unsaturated fatty acids (Volkov *et al.*, 2010).

The first MCRA protein (MW ~ 67 kDa) was identified in *Streptococcus pyogenes*, while screening for antigens recognised by acute rheumatic fever sera. Its amino acid sequence did, at that time, not exhibit similarity to any streptococcal protein with a known function, although it was found to be conserved among pathogenic groups A, C, and G of Streptococci (Kil *et al.*, 1994). The MCRA protein of *Streptococcus pyogenes* M49 was recently shown to represent a flavin adenine dinucleotide (FAD) enzyme which acts as a fatty acid hydratase (Volkov *et al.*, 2010). This latter enzymatic activity catalyses the conversion of OA into 10-hydroxystearic acid, the role of the FAD-binding residues present in this and other (predicted) fatty acid hydratases/MCRA-like proteins are not fully understood (Joo *et al.*, 2012).

MCRA-like proteins show over 50% sequence similarity to the linoleic acid isomerase (LAI) from *Lactobacillus acidophilus* and *Lb. reuteri* PYR8 (Rosson *et al.*, 2001). LAI catalyses the conversion of linoleic acid (LA, a dienoic unsaturated fatty acid) to *cis*-9 *trans*-11 (*c*9,*t*11) conjugated linoleic acid (CLA). CLA is the collective term used to describe a group of polyunsaturated fatty acids that exist as positional and geometric stereoisomer's of octadecadienoic acid (Mulvihill, 2001).

Twenty eight different CLA isomers have been identified, of which the most abundant is the c9,t11 isomer, representing approximately 80 % of total CLA in food products. CLA is found naturally in ruminant food products, for example lamb, beef and dairy, due to the process of biohydrogenation of LA in the rumen (Bhattacharya *et al.*, 2006). The LAI protein is a membrane-bound enzyme which was first isolated from *Butyrivibrio fibrisolvens* (Kepler *et al.*, 1970). Subsequently, the crystal-structure of an LAI from *Propionibacterium acnes* was solved (Liavonchanka *et al.*, 2006) and the reaction mechanism characterised (Liavonchanka *et al.*, 2009).

It was recently suggested (Rosberg-Cody *et al.*, 2011a) that the MCRA-like protein produced by bifidobacteria is responsible for the first step of a two-step CLA production process in which hydroxy-fatty acids act as intermediates in the production of CLA (Ogawa *et al.*, 2001). In the latter study, the gene encoding the MCRA-like protein from *B. breve* NCFB 2258 was cloned, sequenced and expressed in two heterologous hosts (*Lactococcus* and *Corynebacterium*), and the recombinant proteins assessed for enzymatic activity against fatty acid substrates. It was demonstrated that heterologous expression of this bifidobacterial MCRA in *Lactococcus* and *Corynebacterium* resulted in increased amounts of hydroxy-fatty acids in the culture medium. It was also observed that these recombinant hydroxy-fatty acid-producing cells were more resistant to heat and solvent stress as compared to their non-recombinant, wild type controls.

The aim of the current study was to investigate the biological function of the MCRA-like protein in *Bifidobacterium breve*. The obtained results show that the *B. breve* MCRA-like protein functions as an oleate hydratase, which plays a role in stress tolerance in bifidobacteria, though it does not appear to play a role in CLA production in *B. breve*.

2.3 MATERIALS AND METHODS

Bacterial strains, plasmids and culture conditions

Bacterial strains and plasmids used in this study are detailed in (Table 2.1). Bifidobacterial strains were routinely cultured in either de Man Rogosa and Sharpe medium (MRS; Difco™, BD, Le Pont de Claix, France) supplemented with 0.05 % cysteine-HCl, or reinforced clostridial medium (RCM; Oxoid Ltd. Basingstoke, England). Bifidobacterial cultures were incubated at 37^{o}C under anaerobic conditions which were maintained using an anaerocult oxygen depleting system with an atmosphere of 5% CO_2 - 5% H_2 - 90% N_2 (Merck, Darmstadt, Germany) in an anaerobic chamber. *Escherichia coli* strains were cultured in Luria-Bertani broth (LB) (Sambrook *et al.*, 1989) at 37^{o}C with agitation. Where appropriate, growth media contained erythromycin (Em; 100 µg ml^{-1} for *E. coli*), tetracycline (Tet; 10 µg ml^{-1} for *E. coli* and *B. breve)* or kanamycin (Km; 50 µg ml^{-1} for *E. coli*).

Nucleotide sequence analysis

Sequence data were obtained from the Artemis-mediated (Rutherford *et al.*, 2000) genome annotations of the *B. breve* UCC2003 genome (O'Connell Motherway *et al.*, 2011b). Database searches were performed using non-redundant sequences accessible at the National Centre for Biotechnology Information internet site (http://www.ncbi.nlm.nih.gov) using Blast (Altschul *et al.*, 1990; Altschul *et al.*, 1997). Sequence assembly, verification and analysis were performed using the Seqman and Seqbuilder programs of the DNASTAR software package (DNASTAR, Madison, WI, USA_v10.1.2).

DNA manipulations

Chromosomal DNA was isolated from bifidobacteria as previously described (O'Riordan & Fitzgerald, 1999). Minipreparation of plasmid DNA from *E. coli* was achieved using the Qiaprep spin plasmid miniprep kit (Qiagen GmBH, Hilden, Germany). Procedures for DNA manipulations were performed essentially as previously described (Sambrook *et al.,* 1989). Restriction enzymes and T4 DNA ligase were used according to the supplier's instructions (Roche Diagnostics, Bell Lane, East Sussex, UK). Synthetic single stranded oligonucleotide primers used in this study (Supplementary table S2.1) were synthesised by MWG Biotech AG (Ebersberg, Germany). Standard PCRs were performed using TaqPCR mastermix (Qiagen GmBH, Hilden, Germany), *B. breve* colony PCRs were performed using reddymix extensor PCR mastermix (Thermo scientific, Waltham, USA), and both of these PCR procedures were performed according to manufacturer's instructions using a biometra T3000 thermocycler (Biometra, Goettingen, Germany). PCR amplicons were purified using the Qiagen PCR purification kit (Qiagen GmBH, Hilden, Germany). Electroporation of plasmid DNA into *E. coli* or *B. breve* NCFB 2258 was performed as previously described (Maze *et al.*, 2007; Sambrook *et al.*, 1989). The correct orientation and integrity of all constructs was verified by DNA sequencing, performed at MWG Biotech (Ebersberg, Germany).

Construction of B. breve insertion mutant

B. breve NCFB 2258 is highly transformable achieving transformation efficiencies comparable to those achieved with *B. breve* UCC2003 and it was observed that methylation of plasmid DNA in *E. coli* pNZ-MBbrI-MBbrII allows homologous recombination to be achieved in *B. breve* NCFB 2258. An internal fragment of the

putative MCRA-encoding gene *Bbr_1293* (472 bp) was amplified by PCR using *B. breve* UCC2003 chromosomal DNA as template and primer combination MCRAFhd3 and MCRARxba1 (Table 3). *B. breve* UCC2003 sequence was used to design all primers as the *B. breve* NCFB 2258 genome sequence is unavailable. The generated PCR product was ligated to pORI19, an Ori^+ $RepA^-$ integration plasmid (Law *et al.*, 1995), using HindIII and XbaI restriction sites that were incorporated into the primers for the MCRA amplicons and introduced into *E. coli* EC101 by electroporation. Recombinant *E. coli* EC101 derivatives containing pORI19-based constructs were selected on LB agar containing Em, and supplemented with X-gal (5-bromo-4-chloro-3-indolyl-D-galactopyranoside) (40 g ml^{-1}) and 1 mM IPTG.

The expected genetic structure of the recombinant plasmid, designated pORI19-MCRA (pORI19 containing an internal 472 bp fragment of the MCRA-encoding gene), was confirmed by restriction mapping prior to subcloning of the Tet resistance antibiotic cassette, *tetW*, from pAM5 (Alvarez-Martin *et al.*, 2007) as a SacI fragment into the unique SacI site of pORI19-MCRA.

The orientation of the tetracycline (Tet) resistance gene in the resulting plasmid, designated pORI19-tet-MCRA, was determined by restriction analysis. The plasmid was subsequently introduced into *E. coli* EC101 pNZ-MBbrI-MBbrII and transformants were selected based on Cm and Tet resistance. Methylation of the plasmid complement of the obtained transformants in EC101 pNZ -MBbrI-MBbrII was confirmed by their observed resistance to PstI restriction (O'Connell Motherway *et al.*, 2009). The methylated pORI19-tet-MCRA plasmid was introduced into *B. breve* NCFB 2258 by electroporation, with subsequent selection for tetracycline resistance on RCA plates supplemented with Tet. The methylated pORI19-tet-420 plasmid (O'Connell Motherway *et al.*, 2011a) was introduced into *B. breve* NCFB

2258 by electroporation, with subsequent selection for tetracycline resistance on RCA plates supplemented with Tet.

Insertion mutants resulting from site specific homologous recombination were initially confirmed by colony PCR targeting the tetracycline resistance gene *tetW*. This was followed by a second confirmatory PCR adopting a *tetW*-based primer, either forward or reverse depending on the orientation of *tetW*, in combination with a primer specific for each targeted gene to confirm integration at the correct chromosomal position. In this case a product would only be obtained if the correct gene disruption had been achieved.

Rapid analysis method for conjugated linoleic acid production

The ability of isolates to convert free linoleic acid to CLA was assayed by incubating cultures in mMRS (Watson *et al.*, 2013) broth supplemented with free linoleic acid (0.5 mg ml^{-1}) (Sigma-Aldrich, Steinheim, Germany) at 37°C for 72 h, with subsequent assessment of the fatty acid profile using a rapid detection method for CLA production (Barrett *et al.*, 2007). A standard curve (Barrett *et al.*, 2007) demonstrated that an increase in the CLA concentration (from 0 to 0.05 mg ml^{-1}) coincided with a linear increase (R^2 = 0.9985) in absorbance for the c9,t11 CLA isomer up to an absorbance of 2.1. Therefore, the CLA concentrations in culture supernatants with an absorbance at 233 nm less than or equal to 2.1 could be calculated from the linear trend line of the standard curve using the equation $y = 43.431x + 0.0053$.

Gas liquid chromatography (GLC) method for conjugated linoleic acid production

The ability of isolates to convert free linoleic acid to CLA was assayed by incubating cultures in mMRS broth supplemented with free linoleic acid (0.5 mg ml^{-1}) at 37°C for 72 h. Fatty acids were extracted from four grams of the fermented sample following addition of 0.75 mg of the internal standard tridecanoic acid (Sigma-Aldrich, Steinheim, Germany). Two millilitres of isopropanol (Labscan, Bangkok, Thailand) and four millilitres of n-hexane (Labscan, Bangkok, Thailand) were added to the sample followed by vortexing for two minutes. Following centrifugation at 2,197 xg for 5 minutes the resulting clear upper layer was removed to a clean glass tube and evaporated by heating at 45^{o}C under a steady flow of nitrogen. Extracted fatty acids were converted to fatty acid methyl esters (FAMEs) by acid catalysed methylation using 4 % methanolic HCl (Supleco, Pennsylvania, USA) at 60^{o}C for 20 minutes and analysed by GLC analysis (Coakley *et al.*, 2003; Stanton *et al.*, 1997).

Determination of hydratase activity

Determination of hydratase activity was performed as previously published (Rosberg-Cody *et al.*, 2011a), with the following modification: before extraction bacterial cells were ground into a fine powder under liquid nitrogen using a beat mill from Retsch (Haan, Germany) and stored at -80°C. GC/MS-analysis was then performed as indicated in a previous publication (Volkov *et al.*, 2010).

Stress tolerance

B. breve NCFB 2258-MCRA and *B. breve* NCFB 2258-420, which carries a mutation in a gene previously shown to encode a β-galactosidase (control) (O'Connell Motherway *et al.*, 2011a), were grown to OD600 nm 0.4 - 0.5, prior to

stress. Solvent tolerance was determined on addition of ethanol (final concentration 16 %; v/v), after 180 minutes ten microliter volumes of serially diluted samples were spotted in triplicate on reinforced clostridial agar plates. Plates were incubated for 48 hours at 37^{o}C anaerobically. Spots containing between 3 and 30 colony forming units (CFU) were counted and the average was used to calculate the number of CFU per millilitre. The values presented are averages of triplicate experiments

2.4 RESULTS

Analysis of CLA production by various Bifidobacterium species

The ability of various strains of bifidobacteria to convert LA to CLA was analysed using a rapid spectrophotometer-based method (Barrett *et al.*, 2007). Of the 17 bifidobacterial strains tested 11 strains were capable of LA to CLA bioconversion. *B. longum* strains KJOC1, KJOC2, UCC2 and UCC3, and *B. breve* strains UCC1 and UCC2003 did not appear to produce appreciable amounts of CLA from free LA.

Due to the fact that the spectrophotometer-based method merely gives an estimation of CLA content, a more accurate gas liquid chromatography (GLC) method (Coakley *et al.*, 2003) was adopted. Our results demonstrate that *B. breve* NCFB 2258 was the strain with the highest CLA (c9,t11) conversion capability, which was in keeping with the rapid spectrophotometer results in terms of high and low CLA producers (Table 2.2). However, *B. breve* UCC2003 and *B. breve* NCIMB 8807 were shown to exhibit different CLA production capabilities even though *B. breve* UCC2003 is a clonal isolate of *B. breve* NCIMB 8807 (Sheehan *et al.*, 2007). Upon close examination *B. breve* NCIMB 8807 was shown to gradually lose its CLA-producing capability upon repeated sub-cultivation (results not shown). Although the reasons for this unstable CLA-production phenotype are not clear, other CLA-producing strains used in this study did not exhibit such behaviour.

Identification and analysis of a B. breve UCC2003 myosin cross reactive encoding gene / MCRA$_{\text{UCC2003}}$

A gene designated *MCRA*$_{\text{UCC2003}}$, predicted to encode a protein with similarity to other myosin cross reactive antigen (MCRA) proteins was identified from the

annotation of the genome sequence of *B. breve* UCC2003 (O'Connell Motherway *et al.*, 2011b). At protein level, the predicted product of $MCRA_{UCC2003}$ is almost identical (99 % identity) to predicted MCRAs encoded by homologs on the genomes of *Bifidobacterium breve* CETC 7263 (Genbank accession number ADY18551.1), DSM 20213 (Genbank accession number ZP_12586529.1), ACS-071-V-Sch8b (Genbank accession number ZP_06596910.1), while also displaying high levels of identity (80-99%) to annotated MCRA-encoding genes from a range of other bifidobacteria. An exception to this is found for *MCRA* homologs on the genomes of *B. animalis* subsp. *lactis* HN019 (Genbank accession number ZP 02963377) and *B. animalis* subsp. *animalis* ATCC 25527 (Genbank accession number YP 006279649.1), where just 68% identity is observed compare to $MCRA_{UCC2003}$. The apparently monocistronic $MCRA_{UCC2003}$ (1,878 bp) on the *B. breve* UCC2003 genome corresponds to a deduced protein of 625 amino acids (Mw ~ 70.5 kDa), and is located upstream of a ribosomal RNA (rRNA) operon and downstream from a gene predicted to encode a 2,5-diketo-D-gluconic acid reductase A (Fig. 2.1)

A mutation in the MCRA-like gene does not affect CLA production in B. breve NCFB 2258

As mentioned above, MCRA-like proteins exhibit over 50 % sequence similarity to LAI from *Lactobacillus acidophilus* and *Lb. reuteri* PYR8 (Rosson *et al.*, 2001), which have been implicated in the conversion of LA to c9,t11 CLA. It was recently also suggested (Rosberg-Cody *et al.*, 2011a) that these MCRA-like proteins are responsible for the first step of a proposed two-step CLA production process in which hydroxy-fatty acids act as intermediates in the production of CLA (Ogawa *et al.*, 2001). In order to determine if the MCRA-encoding gene is involved in CLA

production, an insertion mutant was created in the MCRA-encoding gene, designated here as $MCRA_{NCFB2258}$, of *B. breve* NCFB 2258 after which the CLA production capability of the mutant strain was analysed. Based on the rapid spectrophotometer method of (Barrett *et al.*, 2007), the *B. breve* NCFB 2258-MCRA insertion mutant was shown to produce the same amount of CLA as the parent strain *B. breve* NCFB 2258, and this result was verified by GLC analysis (Table 2.2). These data convincingly demonstrate that the MCRA-encoding gene does not play a role in CLA production by *B. breve* NCFB 2258, at least under the conditions tested.

Effect of the MCRA insertion mutation on hydratase activity of B. breve NCFB 2258

Bifidobacteria have been shown to exhibit hydratase activity, which allows such bacteria to convert oleic acid into 10-hydroxyoctadecanoic acid otherwise known as 10-hydroxystearic acid (10-HSA) (Rosberg-Cody *et al.*, 2011a). In order to test if hydratase activity was affected in *B. breve* NCFB 2258-MCRA carrying a mutation in $MCRA_{NCFB2258}$, this strain was subjected to a hydratase activity test as previously described (Volkov *et al.*, 2010). In contrast to the control strain *B. breve* NCFB 2258-420, which carries an insertion in *Bbr_0420*, previously characterised as an β-galactosidase (O'Connell Motherway *et al.*, 2011a), the insertion mutant *B. breve* NCFB 2258-MCRA did not exhibit hydratase activity above the threshhold detection level (Fig. 2.2), thus substantiating the hypothesis that the $MCRA_{NCFB2258}$ gene product acts as a hydratase that converts oleic acid into 10-HSA.

MCRA insertion mutation affects the stress tolerance of B. breve NCFB 2258

Heterologous expression of the *B. breve* NCFB 2258 MCRA protein in *Lactococcus* and *Corynebacterium* had previously been shown to cause increased stress tolerance

to 3 % (v/v) butanol (Rosberg-Cody *et al.*, 2011a). These findings suggest that a mutation in this gene in *B. breve* NCFB 2258 would consequently lead to increased sensitivity to certain stressful conditions. In order to test this assumption, *B. breve* NCFB 2258-MCRA and *B. breve* NCFB 2258-420 (control) were grown to an OD600 nm of between 0.4 and 0.5, prior to exposure to solvent stress, which was achieved by the addition of ethanol (final concentration 16 %; v/v). The obtained results show that the *B. breve* NCFB 2258-MCRA insertion mutant is indeed more sensitive to solvent stress than the control strain *B. breve* NCFB 2258-420 in the presence of ethanol (Fig. 2.3).

2.5 DISCUSSION

The aim of the current study was to investigate the biological function of the MCRA-encoding gene of *B. breve* in the hydration of oleic acid into 10-hydroxystearic acid and its possible role in CLA production. We showed that various bifidobacterial strains have different CLA production capabilities with *B. breve* NCFB 2258 and *B. breve* JCM 7017 producing high levels of CLA (Coakley *et al.*, 2003; Hennessy *et al.*, 2009; Rosberg-Cody *et al.*, 2004) while the tested *B. longum* strains did not produce measurable amounts of CLA from LA, a feature common among *B. longum* strains (Ewaschuk *et al.*, 2006; Rosberg-Cody *et al.*, 2004). Interestingly, analysis of the CLA-producing capability of *B. breve* NCIMB 8807 showed that this strain lost this ability during repeated sub-cultivation, a phenomenon that has not been reported previously.

In this study we demonstrated that the MCRA-encoding gene from *B. breve* NCFB 2258 is an oleate hydratase, which is responsible for the breakdown of oleic acid to 10-hydroxystearic acid. MCRA-encoding genes from various bacteria, including bifidobacteria, have been associated with hydratase activity (Bevers *et al.*, 2009; Rosberg-Cody *et al.*, 2011b; Volkov *et al.*, 2010; Yang *et al.*, 2012). The MCRA-like protein has been implicated in the first step of CLA production, an assumption which is based on the finding that CLA production from LA by *L. acidophilus* is presumed to be a two-step process in which hydroxy fatty acids act as intermediates in the production of CLA (Ogawa *et al.*, 2001). This hypothesised mechanism of bacterial CLA production was based on the finding that washed cells of the strain *Lb. acidophilus* AKU 1137 were shown to rapidly convert hydroxy fatty acids to their respective CLA isomers. Therefore CLA production by *Lb. acidophilus* was proposed to involve hydration of LA to 10-hydroxy-18:1, followed by the

dehydration and isomerisation of hydroxy fatty acids to c9,t11 CLA. It is evident in this study that the MCRA-encoding gene of *B. breve* NCFB 2258 is not involved in CLA production and that bifidobacteria follow a different biochemical route to achieve LA to CLA conversion. It is also interesting to note that this *MCRA* gene is present in *B. longum* strains which do not produce CLA such as in this case KJOC1.

In relation to solvent stress, *B. breve* NCFB 2258-MCRA was shown to be more sensitive to 16 % ethanol (v/v) as compared to the control strains *B. breve* NCFB 2258 and *B. breve* NCFB 2258-420. In a previous study (Rosberg-Cody *et al.*, 2011b) increased stress tolerance to butanol was observed when *B. breve* NCFB 2258 MCRA was heterologously expressed in *L. lactis* and *C. glutamicum*. Ethanol was chosen as an alternative membrane stress to butanol as similarities exist in how microbes respond to these membrane stress-invoking compounds (Liu & Qureshi, 2009). The effects of ethanol are a result of dielectric, polar and hydrogen bond interactions with the surface groups of the membrane and membrane proteins, ethanol therefore has the same effects as butanol (Jones, 1989). Previous studies have suggested that MCRA proteins play a role in stress tolerance within the gastrointestinal tract. Deletion of the MCRA-encoding gene in *L. acidophilus* was shown to result in a strain that exhibited reduced growth in the presence of lactate, acetate and salt, which is consistent with a reduced membrane integrity (O'Flaherty & Klaenhammer, 2010).

Ethanol and butanol are short chain alcohols and bacteria respond to these compounds in a similar manner through changes in membrane fatty acid composition, structure and membrane fluidity (Liu & Qureshi, 2009). In Gram negative bacteria exposure to ethanol leads to disruption of cell membrane integrity and structure which leads to cell death. Ethanol-tolerant phenotypes in some *E. coli*

strains are known to result from adaptive changes in the composition of the membrane fatty acids with fatty acid length changing from 16 to 18 carbons, while other changes in the composition of the cell envelope may also confer tolerance (Liu & Qureshi, 2009).

The mechanism underlying Gram positive bacterial tolerance to organic solvents has not been studied extensively (Torres *et al.*, 2011), and have been speculated to be due to the induction of a general stress regulon (Sardessai & Bhosle, 2002), production of organic solvent-emulsifying or deactivating enzymes (Moriya *et al.*, 1995) and/or an active solvent efflux pump (Inoue & Horikoshi, 1991; Moriya *et al.*, 1995). Unsaturated fatty acids are toxic for many bacteria due to their detrimental effect on the cytoplasmic membrane and it has been hypothesised that MCRA ensymes may provide a detoxification mechanism by the hydration of unsaturated fatty acids (Volkov *et al.*, 2010). This detoxification activity may also apply to the bifidobacterial MCRA proteins and may represent a physiological adaptation to the colonic environment.

Collectively our results demonstrate that the MCRA-like protein encoded by *B. breve* NCFB 2258 is an oleate hydratase, which plays a role in stress tolerance in bifidobacteria, while it does not appear to have a role in LA to CLA conversion by *B. breve.* The actual metabolic pathway of CLA production in bifidobacteria is thus still elusive and our future research efforts will aim to uncover the genetic elements involved in this bioconversion.

2.6 ACKNOWLEDGEMENTS

The Alimentary Pharmabiotic Centre is funded by Science Foundation Ireland (SFI), through the Irish Government's National Development Plan. The authors and their work were supported by SFI (grant no's 02/CE/B124 and 07/CE/B1368), the Tomar trust, which supported a postgraduate fellowship to KJO'C, and a HRB postdoctoral fellowship (Grant no. PDTM/20011/9) awarded to MOCM. Technical assistance from Sabine Freitag and help with GC/MS-analysis from Dr. Cornelia Herrfurth is gratefully acknowledged.

2.7 TABLES AND FIGURES

Table 2.1 Bacterial strains and plasmids used in this study

Strain/ plasmid	Relevant characteristics	Reference or source
E.coli strains		
EC101-pNZ-M.Bbrll+Bbr111	EC101 harbouring pNZ8048 derivative containing bbrllM and bbrlllM	(MaryO'Connell-Motherway *et al.*, 2009)
EC101	Cloning host, $repA^+$ km^r	(Law *et al.*, 1995)
B.breve strains		
UCC2003	Isolate from a nursling stool	(Maze *et al.*, 2007)
NCIMB 8807	Isolate from a nursling stool	NCIMB
NCFB 2257	Isolate from infant intestine	NCFB
NCTC 11815	Isolate from infant intestine	NCTC
NCFB2258	Isolate from infant intestine	NCFB
NCIMB 8815	Isolate from infant faeces	NCIMB
JCM 7017	Isolate from human faeces	JCM
UCC2005	Isolate from human faeces	UCC
UCC2007	Isolate from human faeces	UCC
NIZO658	Isolate from a nursling stool	NIZO
LMG 13208	Isolate from infant intestine	LMG
UCC1	Isolate from human faeces	UCC
NCFB2258-MCRA	pORI19-tet-MCRA insertion mutant of 2258	This study
NCFB2258-420	pORI19-tet-420 insertion mutant of 2258	This study
B.longum Strains		UCC
UCC2	Isolate from human faeces	UCC
UCC3	Isolate from human faeces	UCC
KJOC1	Isolate from infant faeces	
KJOC2	Isolate from infant faeces	
Plasmids		
pORI19	Em^r, $repA^-$, ori^+, cloning vector	(Law *et al.*, 1995)
pAM5	pBC1-puC19-Tc^r	(Alvarez-Martin*et al.*, 2007)

JCM: Japan Collection of Microorganisms; NIZO: Nizo Food Research; LMG: Belgian coordinated Collection of Microorganisms; NCFB: National Collection of Food Bacteria; NCIMB: National Collection of Industrial and Marine Bacteria; NCTC: National Collection of Type Cultures; UCC: University College Cork Culture Collection.

Table 2.2 Bioconversion of linoleic acid to CLA by bifidobacterial strains

			CLA converted from 0.5mg ml^{-1} linoleic acid*		
Strain	**Species**	**Source**	***c*9,*t*11**	**S.D**	**% Converted**
Bifidobacterium breve	UCC2003	Isolate from a nursling stool	0.008	0.001	1.73
	NCIMB 8807	Isolate from a nursling stool	0.085	0.011	17.16
	NCFB 2257	Isolate from infant intestine	0.027	0.006	5.54
	NCTC 11815	Isolate from infant intestine	0.148	0.011	29.64
	NCFB 2258	Isolate from infant intestine	0.245	0.021	49.0
	NCIMB 8815	Isolate from infant faeces	0.077	0.021	15.54
	JCM 7017	Isolate from human faeces	0.229	0.027	45.90
	UCC2005	Isolate from human faeces	0.179	0.033	35.88
	UCC2007	Isolate from human faeces	0.054	0.003	10.84
	Nizo 658	Isolate from a nursling stool	0.083	0.004	16.70
	LMG 13208	Isolate from infant intestine	0.055	0.00	11.16
	UCC1	Isolate from human faeces	0.005	0.000	1.05
	NCFB 2258-MCRA	This study	0.213	0.002	42.60
Bifidobacterium longum					
	UCC2	Isolate from human faeces	0	0	0
	UCC3	Isolate from human faeces	0	0	0
	KJOC1	Isolate from infant faeces	0	0	0
	KJOC2	Isolate from infant faeces	0	0	0

* Values represent the average of two independent experiments

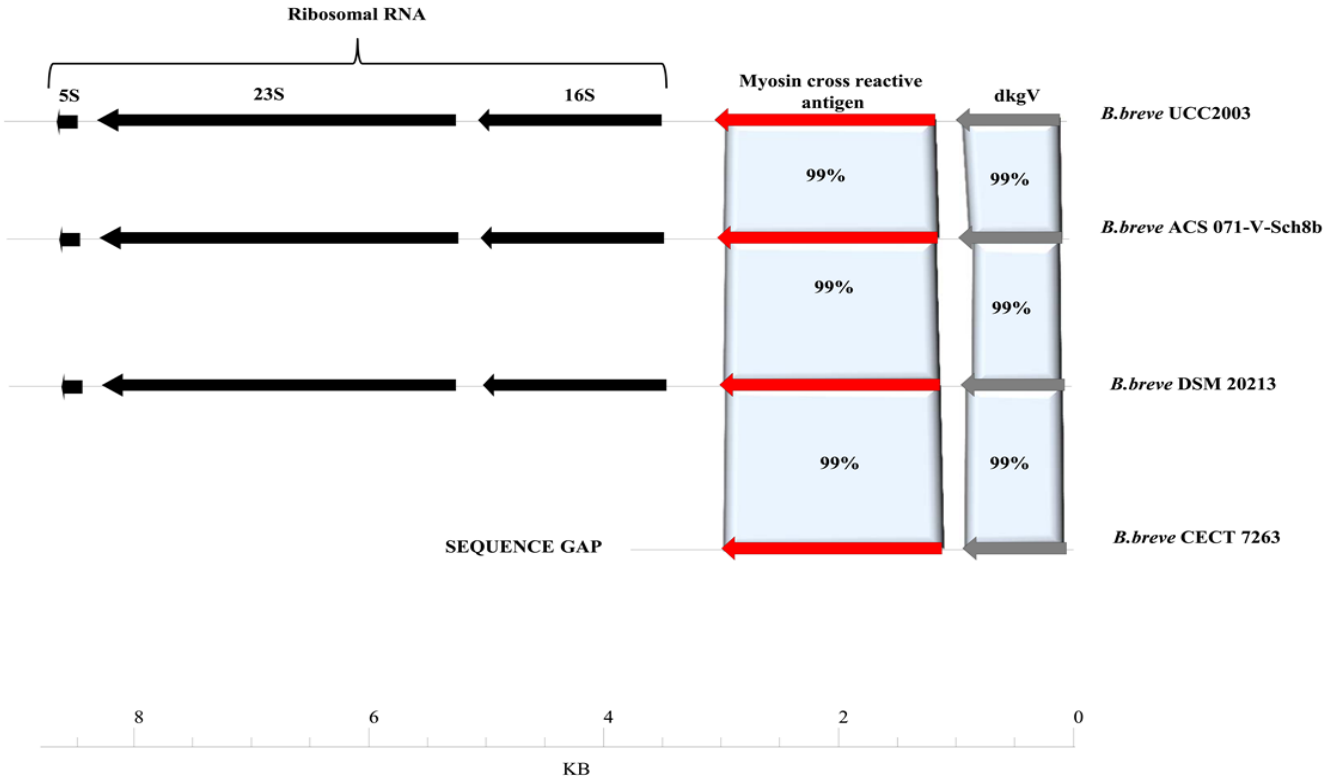

Figure 2.1 Schematic representation of the comparison of the myosin cross reactive antigen encoding gene, $MCRA_{UCC2003}$, from *B. breve* UCC2003 to other sequenced *B. breve* (O'Connell Motherway *et al.*, 2009).

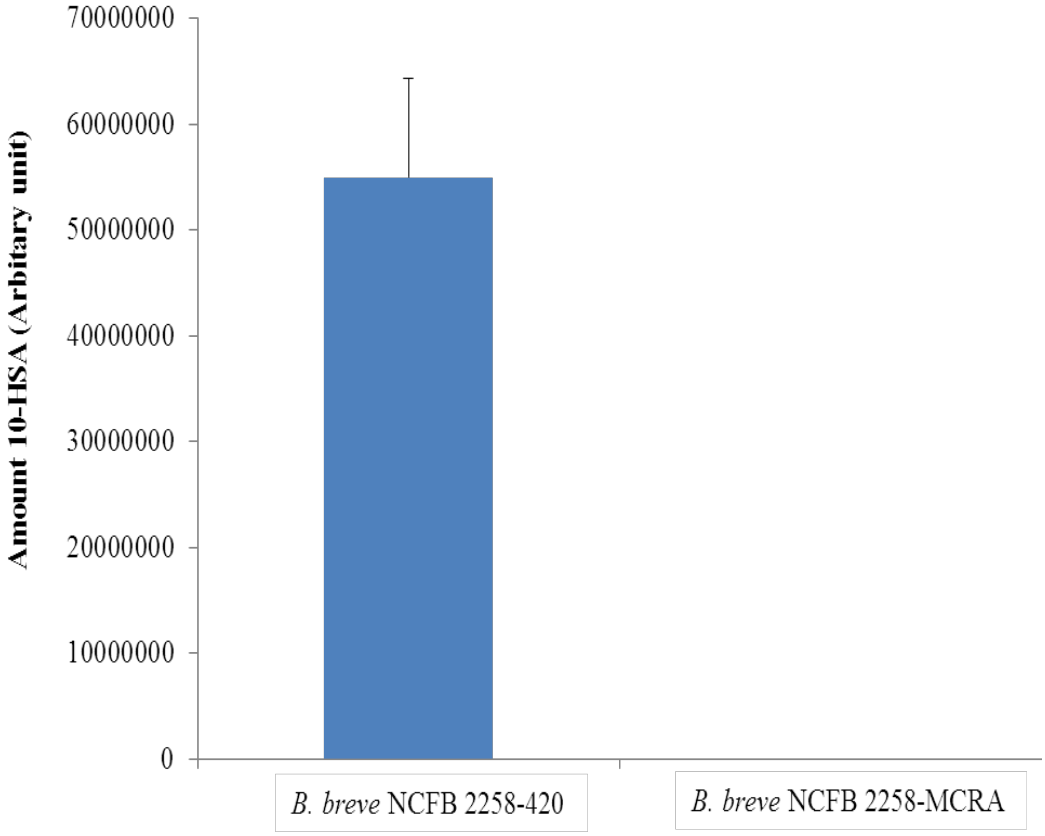

Figure 2.2 The ability of *B. breve* NCFB 2258-MCRA and *B. breve* NCFB 2258-420 (control) to convert oleic acid to 10-hydroxystearic acid was assayed by incubating cultures in MRS broth which contains 1% oleic acid at 37°C for 16 h, with subsequent assessment of the fatty acid profile.

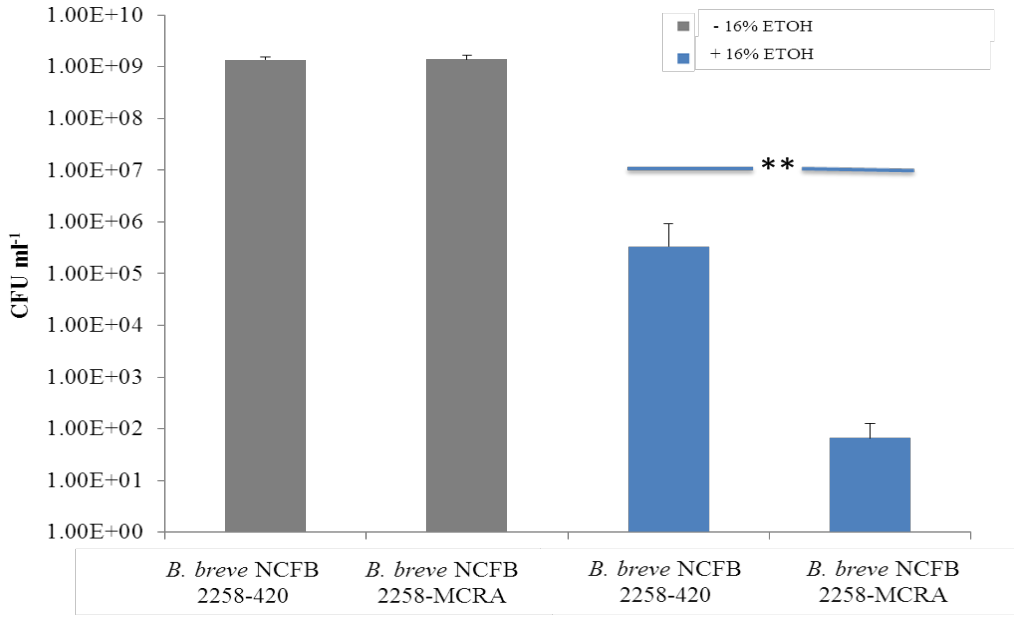

Figure 2.3 *B. breve* NCFB 2258-MCRA and *B. breve* NCFB 2258-420 (control) grown in the absence and presence of ethanol 16% (v/v). Log phase cells were grown to an OD600 nm of 0.4 - 0.5, prior to stress of ethanol 16% (v/v), after 180 minutes cultures were spot plated, incubated at 37^{o}C for 48hrs anaerobically this was followed by viable cell counts. The values represent the average of three independent experiments with standard error. *P <0.001

2.8 SUPPLEMENTARY MATERIAL

Table S2.1 Oligonucleotide primers used in this study

Purpose	**Primer**	**Sequence**[a]	**Size**
Amplification of MCRA internal fragment from *B. breve* UCC2003 to create mutant	MCRAFhd3 MCRARxba1	TGCATC**AAGCTT**CGATGACGAGGTGCTGAAC TGCGCA**TCTAGA**CAGCCGCCGTTGGTGATG	472bp
Confirmation of MCRA integration in *B. breve* NCFB 2258	MCRA -confirm Tetwsal1F	CTACAGCAGCGGCAACTATG TCAGCT**GTCGAC**ATGCTCATGTACGGTAAG	3.5KB
Confirmation of 420 integration in *B. breve* NCFB 2258	420-confirm Tetsal1F	GTTGCTACTCCCTCTGACTCTCC TCAGCT**GTCGAC**ATGCTCATGTACGGTAAG	3.2KB 2.2KB
Confirmation of *tetw* integration in *B. breve* NCFB 2258	Tetwsal1F Tetwsal1R	TCAGCT**GTCGAC**ATGCTCATGTACGGTAAG GCGACG**GTCGAC**CATTACCTTCTGAAACATA	

[a] Sequences of restriction enzyme sites are indicated in bold

2.9 REFERENCES

Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. (1990). Basic local alignment search tool. *Journal of Molecular Biology* **215**, 403-410.

Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J., Zhang, Z., Miller, W. & Lipman, D. J. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Research* **25**, 3389-3402.

Alvarez-Martin, P., O'Connell Motherway, M., van Sinderen, D. & Mayo, B. (2007). Functional analysis of the pBC1 replicon from *Bifidobacterium catenulatum* L48. *Applied Microbiology and Biotechnology* **76**, 1395-1402.

Barrett, E., Ross, R. P., Fitzgerald, G. F. & Stanton, C. (2007). Rapid Screening Method for Analysing the Conjugated Linoleic Acid Production Capabilities of Bacterial Cultures. *Applied and Environmental Microbiology* **73**, 2333-2337.

Bevers, L. E., Pinkse, M. W. H., Verhaert, P. D. E. M. & Hagen, W. R. (2009). Oleate Hydratase Catalyses the Hydration of a Nonactivated Carbon-Carbon Bond. *Journal of Bacteriology* **191**, 5010-5012.

Bhattacharya, A., Banu, J., Rahman, M., Causey, J. & Fernandes, G. (2006). Biological effects of conjugated linoleic acids in health and disease. *The Journal of Nutritional Biochemistry* **17**, 789-810.

Coakley, M., Ross, R. P., Nordgren, M., Fitzgerald, G., Devery, R. & Stanton, C. (2003). Conjugated linoleic acid biosynthesis by human-derived *Bifidobacterium* species. *Journal of Applied Microbiology* **94**, 138-145.

Ewaschuk, J. B., Walker, J. W., Diaz, H. & Madsen, K. L. (2006). Bioproduction of conjugated linoleic acid by probiotic bacteria occurs in vitro and in vivo in mice. *The Journal of Nutrition* **136**, 1483-1487.

Greenway, D. L. & Dyke, K. G. (1979). Mechanism of the inhibitory action of linoleic acid on the growth of *Staphylococcus aureus*. *Journal of General Microbiology* **115**, 233-245.

Hennessy, A. A., Ross, R. P., Devery, R. & Stanton, C. (2009). Optimisation of a reconstituted skim milk based medium for enhanced CLA production by bifidobacteria. *Journal of Applied Microbiology* **106**, 1315-1327.

Inoue, A. & Horikoshi, K. (1991). Estimation of solvent-tolerance of bacteria by the solvent parameter log P. *Journal of Fermentation and Bioengineering* **71**, 194-196.

Jones, R. P. (1989). Biological principles for the effects of ethanol. *Enzyme and Microbial Technology* **11**, 130-153.

Joo, Y. C., Seo, E. S., Kim, Y. S., Kim, K. R., Park, J. B. & Oh, D. K. (2012). Production of 10-hydroxystearic acid from oleic acid by whole cells of recombinant *Escherichia coli* containing oleate hydratase from *Stenotrophomonas maltophilia*. *Journal of Biotechnology* **158**, 17-23.

Kepler, C. R., Tucker, W. P. & Tove, S. B. (1970). Biohydrogenation of Unsaturated Fatty Acids: iv. substrate specificity and inhibition of linoleate δ12-cis, δ11-trans-isomerase from *Butyrivibrio fibrisolvens*. *Journal of Biological Chemistry* **245**, 3612-3620.

Kil, K. S., Cunningham, M. W. & Barnett, L. A. (1994). Cloning and sequence analysis of a gene encoding a 67-kilodalton myosin-cross-reactive antigen of *Streptococcus pyogenes* reveals its similarity with class II major histocompatibility antigens. *Infection and Immunity* **62**, 2440-2449.

Law, J., Buist, G., Haandrikman, A., Kok, J., Venema, G. & Leenhouts, K. (1995). A system to generate chromosomal mutations in *Lactococcus lactis* which allows fast analysis of targeted genes. *Journal of Bacteriology* **177**, 7011-7018.

Leahy, S. C., Higgins, D. G., Fitzgerald, G. F. & van Sinderen, D. (2005). Getting better with bifidobacteria. *Journal of Applied Microbiology* **98**, 1303-1315.

Liavonchanka, A., Hornung, E., Feussner, I. & Rudolph, M. (2006). In-house SIRAS phasing of the polyunsaturated fatty-acid isomerase from *Propionibacterium acnes. Acta crystallographica Section F, Structural Biology and Crystallization Communications* **62**, 153-156.

Liavonchanka, A., Rudolph, M. G., Tittmann, K., Hamberg, M. & Feussner, I. (2009). On the mechanism of a polyunsaturated fatty acid double bond isomerase from *Propionibacterium acnes. The Journal of Biological Chemistry* **284**, 8005-8012.

Liu, S. & Qureshi, N. (2009). How microbes tolerate ethanol and butanol. *New Biotechnology* **26**, 117-121.

Liu, Z., Jiang, Z., Zhou, K., Li, P., Liu, G. & Zhang, B. (2007). Screening of bifidobacteria with acquired tolerance to human gastrointestinal tract. *Anaerobe* **13**, 215-219.

Maze, A., O'Connell Motherway, M., Fitzgerald, G. F., Deutscher, J. & van Sinderen, D. (2007). Identification and characterisation of a fructose phosphotransferase system in *Bifidobacterium breve* UCC2003. *Applied and Environmental Microbiology* **73**, 545-553.

Moriya, K., Yanigitani, S., Usami, R. & Horikoshi, K. (1995). Isolation and some properties of an organic solvent marine bacterium degrading cholesterol. *Journal of Marine Biotechnology* **2**, 131-133.

Mulvihill, B. (2001). Ruminant meat as a source of conjugated linoleic acid (CLA). *Nutrition Bulletin* **26**, 295-299.

O'Connell, K. J., O'Connell Motherway, M., Hennessey, A. A., Brodhun, F., Ross, R. P., Feussner, I., Stanton, C., Fitzgerald, G. F. & van Sinderen, D. (2013). Identification and characterisation of an oleate hydratase-encoding gene from *Bifidobacterium breve. Bioengineered* **4**, 313–321.

O'Connell Motherway, M., O'Driscoll, J., Fitzgerald, G. F. & van Sinderen, D. (2009). Overcoming the restriction barrier to plasmid transformation and targeted mutagenesis in *Bifidobacterium breve* UCC2003. *Microbial Biotechnology* **2**, 321-332.

O'Connell Motherway, M., Fitzgerald, G. F. & van Sinderen, D. (2011a). Metabolism of a plant derived galactose-containing polysaccharide by *Bifidobacterium breve* UCC2003. *Microbial Biotechnology* **4**, 403-416.

O'Connell Motherway, M., Zomer, A., Leahy, S. C., Reunanen, J., Bottacini, F., Claesson, M.J., O'Brien, F., Flynn, K., Casey, P.G., Munoz, J.A., Kearney, B.,

Houston, A.M., O' Mahony, C., Higgins, DG., Shanahan, F., Palva, A., de Vos, W.M., Fitzgerald, G.F., Ventura, M., O'Toole, P.W. & van Sinderen D. (2011b). Functional genome analysis of *Bifidobacterium breve* UCC2003 reveals type IVb tight adherence (Tad) pili as an essential and conserved host-colonisation factor. *Proceedings of the National Academy of Sciences of the United States of America* **108**, 11217-11222.

O'Flaherty, S. J. & Klaenhammer, T. R. (2010). Functional and phenotypic characterisation of a protein from *Lactobacillus acidophilus* involved in cell morphology, stress tolerance and adherence to intestinal cells. *Microbiology* **156**, 3360-3367.

O'Riordan, K. & Fitzgerald, G. F. (1999). Molecular characterisation of a 5.75-kb cryptic plasmid from *Bifidobacterium breve* NCFB 2258 and determination of mode of replication. *FEMS Microbiology Letters* **174**, 285-294.

Ogawa, J., Matsumura, K., Kishino, S., Omura, Y. & Shimizu, S. (2001). Conjugated linoleic acid accumulation via 10-hydroxy-12-octadecaenoic acid during microaerobic transformation of linoleic acid by *Lactobacillus acidophilus*. *Applied and Environmental Microbiology* **67**, 1246-1252.

Perusi, C., Gomez-Lira, M., Mottes, M. Pignatti, P.F., Bertini, E., Cappa, M., Vigliani, M.C., Schiffer, D., Rizzuto, N. & Salviati A. (1999). Two novel missense

mutations causing adrenoleukodystrophy in Italian patients. *Molecular and Cellular Probes* **13**, 179-182.

Raychowdhury, M. K., Goswami, R. & Chakrabarti, P. (1985). Effect of unsaturated fatty acids in growth inhibition of some penicillin-resistant and sensitive bacteria. *The Journal of Applied Bacteriology* **59**, 183-188.

Rosberg-Cody, E., Ross, R. P., Hussey, S., Ryan, C. A., Murphy, B. P., Fitzgerald, G. F., Devery, R. & Stanton, C. (2004). Mining the microbiota of the neonatal gastrointestinal tract for conjugated linoleic acid-producing bifidobacteria. *Applied and Environmental Microbiology* **70**, 4635-4641.

Rosberg-Cody, E., Liavonchanka, A., Gobel, C., Ross, R. P., O'Sullivan, O., Fitzgerald, G. F., Feussner, I. & Stanton, C. (2011a). Myosin-cross-reactive antigen (MCRA) protein from *Bifidobacterium breve* is a FAD-dependent fatty acid hydratase which has a function in stress protection. *BMC Biochemistry* **12**, 9.

Rosberg-Cody, E., Stanton, C., O'Mahony, L., Wall, R., Shanahan, F., Quigley, E. M., Fitzgerald, G. F. & Ross, R. P. (2011b). Recombinant *lactobacilli* expressing linoleic acid isomerase can modulate the fatty acid composition of host adipose tissue in mice. *Microbiology* **157**, 609-615.

Rosson, R. A., Grund, A., Deng, M. & Sanchez-Riera, F. (2001). Linoleate isomerase. *World Patent* **100846**, 30.

Rutherford, K., Parkhill, J., Crook, J., Horsnell, T., Rice, P., Rajandream, M. A. & Barrell, B. (2000). Artemis: sequence visualisation and annotation. *Bioinformatics* **16**, 944-945.

Sambrook, J., Fritsch, E. F. & Maniatis, T. (1989). *Molecular cloning: a laboratory manual*: Cold Spring Harbour Laboratory.

Sardessai, Y. & Bhosle, S. (2002). Tolerance of bacteria to organic solvents. *Research in Microbiology* **153**, 263-268.

Sheehan, V. M., Sleator, R. D., Hill, C. & Fitzgerald, G. F. (2007). Improving gastric transit, gastrointestinal persistence and therapeutic efficacy of the probiotic strain *Bifidobacterium breve* UCC2003. *Microbiology* **153**, 3563-3571.

Stanton, C., Lawless, F., Kjellmer, G., Harrington, D., Devery, R., Connolly, J. F. & Murphy, J. (1997). Dietary Influences on Bovine Milk cis-9,trans-11-Conjugated Linoleic Acid Content. *Journal of Food Science* **62**, 1083-1086.

Stanton, C., Ross, R. P., Fitzgerald, G. F. & Sinderen, D. V. (2005). Fermented functional foods based on probiotics and their biogenic metabolites. *Current Opinion in Biotechnology* **16**, 198-203.

Tannock, G. W. (1997). Probiotic properties of lactic-acid bacteria: plenty of scope for fundamental R & D. *Trends in Biotechnology* **15**, 270-274.

Terés, S., Barceló-Coblijn, G., Benet, M., Álvarez, R., Bressani, R., Halver, J. E. & Escribá, P. V. (2008). Oleic acid content is responsible for the reduction in blood pressure induced by olive oil. *Proceedings of the National Academy of Sciences* **105**, 13811-13816.

Tissier, H. (1900). Recherchers sur la flora intestinale normale pathologique du nourisson. *Thesis, University of Paris, Paris France*.

Torres, S., Pandey, A. & Castro, G. R. (2011). Organic solvent adaptation of Gram positive bacteria: applications and biotechnological potentials. *Biotechnology Advances* **29**, 442-452.

Turroni, F., Ribbera, A., Foroni, E., van Sinderen, D. & Ventura, M. (2008). Human gut microbiota and bifidobacteria: from composition to functionality. *Antonie van Leeuwenhoek* **94**, 35-50.

Turroni, F., Peano, C., Pass, D. A., Foroni, E., Severgnini, M., Claesson, M. J., Kerr, C., Hourihane, J., Murray, D., Fuligni, F., Gueimonde, M., Margolles, A., De Bellis, G., O'Toole, P. W., van Sinderen, D., Marchesi, J. R., & Ventura, M. (2012). Diversity of Bifidobacteria within the Infant Gut Microbiota. *PLoS ONE* **7**, e36957.

Ventura, M., van Sinderen, D., Fitzgerald, G. F. & Zink, R. (2004). Insights into the taxonomy, genetics and physiology of bifidobacteria. *Antonie van Leeuwenhoek* **86**, 205-223.

Ventura, M., O'Connell Motherway, M., Leahy, S., Moreno-Munoz, J. A., Fitzgerald, G. F. & van Sinderen, D. (2007). From bacterial genome to functionality; case bifidobacteria. *International Journal of Food Microbiology* **120**, 2-12.

Ventura, M., Turroni, F., O'Connell Motherway, M., MacSharry, J. & van Sinderen, D. (2012). Host-microbe interactions that facilitate gut colonisation by commensal bifidobacteria. *Trends in Microbiology* **20**, 467-476.

Volkov, A., Liavonchanka, A., Kamneva, O., Fiedler, T., Goebel, C., Kreikemeyer, B. & Feussner, I. (2010). Myosin cross-reactive antigen of *Streptococcus pyogenes* M49 encodes a fatty acid double bond hydratase that plays a role in oleic acid detoxification and bacterial virulence. *The Journal of Biological Chemistry* **285**, 10353-10361.

Wallen, L. L., Benedict, R. G. & Jackson, R. W. (1962). The microbiological production of 10-Hydroxystearic acid from oleic acid. *Archives of Biochemistry and Biophysics* **99**, 249-253.

Watson, D., O'Connell Motherway, M., Schoterman, M. H. C., van Neerven, R. J. J., Nauta, A. & van Sinderen, D. (2013). Selective carbohydrate utilisation by *lactobacilli* and bifidobacteria. *Journal of Applied Microbiology*, **114,** 1132-1146.

Yang, B., Chen, H., Song, Y., Chen, Y. Q., Zhang, H. & Chen, W. (2012). Myosin-cross-reactive antigens from four different lactic acid bacteria are fatty acid hydratases. *Biotechnology Letters*, **35**, 75-81.

Zheng, C. J., Yoo, J. S., Lee, T. G., Cho, H. Y., Kim, Y. H. & Kim, W. G. (2005). Fatty acid synthesis is a target for antibacterial activity of unsaturated fatty acids. *FEBS Letters* **579**, 5157-5162.

Chapter III

Metabolism of four α-glycosidic linkage-containing oligosaccharides by *Bifidobacterium breve* UCC2003

Schematic of *Bifidobacterium* genome comparison was provided by Francesca Bottacini.

This chapter was published in:

O'Connell, K. J., O'Connell Motherway, M., O'Callaghan, J., Fitzgerald, G. F., Ross, R. P., Ventura, M., Stanton, C. & van Sinderen, D. (2013). Metabolism of four α-glycosidic linkage-containing oligosaccharides by *Bifidobacterium breve* UCC2003. *Applied and Environmental Microbiology* **79**, 6280-6292

3.1 ABSTRACT

Members of the genus *Bifidobacterium* are common inhabitants of the gastrointestinal tract of humans and other mammals, where they ferment many diet-derived carbohydrates that cannot be digested by their host. To extend our understanding of bifidobacterial carbohydrate utilisation, we investigated the molecular mechanisms by which eleven strains of *Bifidobacterium breve* metabolise four distinct α-glucose and/or α-galactose-containing oligosaccharides. Here we demonstrate that all *B. breve* strains examined possess the ability to utilise raffinose, stachyose and melibiose. However, the ability to metabolise melezitose was not ubiquitous for all tested *B. breve* strains. Transcriptomic and functional genomic approaches identified a gene cluster dedicated to the metabolism of α-galactose-containing carbohydrates, while an adjacent gene cluster, dedicated to the metabolism of α-glucose-containing melezitose, was identified in strains able to use this carbohydrate.

3.2 INTRODUCTION

Bifidobacteria are Gram positive, saccharolytic, non-motile, non-sporulating, anaerobic rods which possess a high GC genome content, and belong to the phylum *Actinobacteria* and the family *Bifidobacteriaceae* (Miyake *et al.*, 1998; Ventura *et al.*, 2007b). The first representatives of the genus *Bifidobacterium* were isolated more than a century ago (Tissier, 1900), and these bacteria naturally inhabit the gastrointestinal tract of humans and other mammals, being particularly abundant in breast-fed infants (Turroni *et al.*, 2008; Turroni *et al.*, 2012a). Certain bifidobacterial strains have been claimed to promote and maintain gastrointestinal health (Tannock, 1997) and are therefore used as health-promoting or probiotic bacterial ingredients in certain functional foods (Stanton *et al.*, 2005). Their reported beneficial effects on the host include, among others, inhibition of bacterial and viral pathogens, alleviating lactose intolerance, enhancing natural immunity and reducing serum cholesterol (Leahy *et al.*, 2005; Liu *et al.*, 2007; Ventura *et al.*, 2012).

As saccharolytic microorganisms, bifidobacteria degrade various oligosaccharides and polysaccharides into their monosaccharide constituents, which are then shuttled into a specific hexose fermentation pathway called the fructose-6-phosphate phosphoketolase pathway or bifid-shunt (Scardovi & Trovatelli, 1965). Species- or strain-specific proliferation of commensal bifidobacteria is presumed to be stimulated by particular dietary carbohydrates, which are for this reason termed prebiotic substances (Macfarlane *et al.*, 2008). The term prebiotic was first coined in 1995 (Gibson & Roberfroid, 1995), when it was defined as “non-digestible food ingredients that beneficially affect the host by selectively stimulating the growth and/or activity of one or a limited number of bacteria in the colon”. Up to 8 % of the coding capacity of a bifidobacterial genome is dedicated to carbohydrate

metabolism, of which half are considered to be responsible for carbohydrate uptake, mainly via ATP binding cassette (ABC) transporters, although proton motive force-driven permeases, proton symporters, and phosphoenolpyruvate-dependent phosphotransferase systems (PEP-PTSs) may also be employed for this purpose (Maze *et al.*, 2007; Pokusaeva *et al.*, 2011a; Schell *et al.*, 2002). A range of glycosyl hydrolases, enzymes that hydrolyse a specific glycosidic bond between the monosaccharidic moieties of certain oligo- and polysaccharides, allow bifidobacteria to grow on dietary and host-derived carbohydrates present in the gastrointestinal tract (Pokusaeva *et al.*, 2011a).

Certain bifidobacterial strains have previously been shown to grow on soymilk-derived α-galacto-oligosaccharides, such as raffinose [α-D-Gal*p*-(1→6)-α-D-Glc*p*-(1→2)-β-D-Fru*f*], stachyose [α-D-Gal*p*-(1→6)-α-D-Gal*p*-(1→6)-α-D-Glc*p*-(1→2)-β-D-Fru*f*] and melibiose [α-D-Gal*p*-(1→6)-α-D-Glc*p*] (Garro *et al.*, 1999; Minami *et al.*, 1983; Yazawa *et al.*, 1978). Stachyose and raffinose (sugars of the so-called raffinose family, which also includes verbascose) are present in a wide variety of plants (French, 1954), while the related sugar melibiose (though not a member of the raffinose family) is also found in many plants, being particularly abundant in soybean roots and stems (Rehms & Barz, 1995). To metabolise such α-galacto-oligosaccharides bifidobacteria require α-galactosidase enzyme activity, which has been identified and characterised in five bifidobacterial species, *B.bifidum* JCM 1254 (Wakinaka *et al.*, 2013) *B. adolescentis* (Leder *et al.*, 1999; Van Laere *et al.*, 1999), *B. bifidum* (Goulas *et al.*, 2009), *B. breve* (Zhao *et al.*, 2008) and *B. longum* subsp. *longum* (Garro *et al.*, 1994; Hirayama *et al.*, 2012).

Raffinose utilisation and its transcriptional regulation have been characterised in more detail in *Escherichia coli* than in bifidobacteria. In *E. coli,* raffinose is actively transported into the cell using a dedicated raffinose permease (encoded by *rafA*), after which raffinose is hydrolysed into sucrose and galactose by an α-galactosidase (specified by *rafB*), followed by hydrolysis of the sucrose into glucose and fructose by a sucrose hydrolase, which is encoded by *rafD* (Aslanidis *et al.*, 1989).

Various α-glucosidases are produced by *Bifidobacterium* sp. (Pokusaeva *et al.*, 2009; van den Broek *et al.*, 2003), such as the enzymes encoded by *agl1* and *agl2*, which were previously identified and characterised in *B. breve* UCC2003 (Pokusaeva *et al.*, 2009). These two enzymes, both members of glycosyl hydrolase (GH) family 13 which mainly represents enzymes with α-(1→6)-glucosidase activity (EC 3.2.1.10), were shown to exhibit hydrolytic activity towards panose, isomaltose, isomaltotriose, as well as four sucrose isomers, palatinose, trehalulose, turanose and maltulose, while they were also shown to partially degrade trehalose and nigerose. The preferred substrates for the Agl1 and Agl2 enzymes were shown to be panose, isomaltose, and trehalulose, carbohydrates that contain either an α-(1→6)-glucosidic bond (present in panose and isomaltose) or an α-(1→1)-glucosidic bond (present in trehalulose) (Pokusaeva *et al.*, 2009).

Melezitose [α-D-Glc*p*-(1→2)-β-D-Fru*f*-(3→1)-α-D-Glc*p*] is an α-glucose-containing trisaccharide found in honeydew and manna, which are sugar-rich liquid and solid deposits, respectively, associated with leaves and branches of various trees and shrubs (Bacon & Dickinson, 1957). Although it was initially believed that melezitose is an oligosaccharide that is naturally present in various plants (Hudson, 1946), it was later concluded that certain insects are responsible for melezitose production as this sugar is absent from the tree sap used by such insects to form honeydew (Bacon

& Dickinson, 1957). To the best of our knowledge no information is available on how (bifido)bacteria metabolise melezitose.

In the current study, we describe the identification of two adjacent gene clusters, *mel* and *raf*, present in the genome of *B. breve* UCC2003 which are involved in the metabolism of melezitose and raffinose-family/related sugars, respectively.

3.3 MATERIALS AND METHODS

Bacterial strains, plasmids and culture conditions

Bacterial strains and plasmids used in this study are detailed in (Table 3.1). Bifidobacteria were routinely cultured in either de Mann Rogosa and Sharpe medium (MRS; Difco™, BD, Le Pont de Claix, France), supplemented with 0.05 % cysteine-HCl, or reinforced clostridial medium (RCM; Oxoid Ltd. Basingstoke, England). Carbohydrate utilisation by bifidobacterial strains was examined in modified de Mann Rogosa and Sharpe Medium (mMRS) prepared from first principles (De Man *et al.*, 1960), devoid of a carbohydrate source. Prior to inoculation, the mMRS medium was supplemented with cysteine-HCl (0.05 % w/v) and a particular carbohydrate source (1 % w/v). The carbohydrates used were raffinose, stachyose, melezitose, melibiose and glucose (all purchased from Sigma-Aldrich, Steinheim, Germany). Bifidobacterial cultures were incubated at 37^{o}C under anaerobic conditions which were maintained using an Anaerocult oxygen depleting system (Merck, Darmstadt, Germany) in an anaerobic chamber. *Lactococcus lactis* strains were cultivated in M17 broth containing 0.5 % glucose (Terzaghi & Sandine, 1975) at 30^{o}C. *E. coli* strains were cultured in Luria-Bertani broth (LB) (Sambrook *et al.*, 1989) at 37^{o}C with agitation. Where appropriate, growth media contained chloramphenicol (Cm; 5µg ml^{-1} for *L. lactis*, 10 µg ml^{-1} for *E. coli*, and 2.5 µg ml^{-1} for *B. breve*), erythromycin (Em; 100 µg ml^{-1} for *E. coli*), tetracycline (Tet; 10 µg ml^{-1} for *E. coli* or *B. breve)* or kanamycin (Km; 50 µg ml^{-1} for *E. coli*).

In order to determine bacterial growth profiles and final optical densities, 5 ml of freshly prepared mMRS medium including a particular carbohydrate (see above) was inoculated with 50 µl (1 %) of a stationary phase culture of a particular *B. breve* strain. Un-inoculated mMRS was used as a negative control. Cultures were

incubated anaerobically at 37°C for 16 hours, and the optical density at 600 nm (OD_{600}) was determined during this period at 30 minute intervals using a Powerwave™ microplate spectrophotometer (BioTek Instruments, Inc. USA) in conjunction with Gen5™ Microplate Software for Windows.

Nucleotide sequence analysis

Sequence data was obtained from the Artemis-mediated (Rutherford *et al.*, 2000) genome annotations of *B. breve* UCC2003 (O'Connell Motherway *et al.*, 2011b). Database searches were performed using non-redundant sequences accessible at the National Centre for Biotechnology Information (NCBI) (http://www.ncbi.nlm.nih.gov) using the basic local alignment search tool (BLAST) (Altschul *et al.*, 1990; Altschul *et al.*, 1997). Sequence verification and analysis were performed using the Seqman and Seqbuilder programs of the DNASTAR software package (DNASTAR, Madison, WI, USA v10.1.2).

DNA manipulations

Chromosomal DNA was isolated as previously described (O'Riordan & Fitzgerald, 1999). Minipreparation of plasmid DNA from *E. coli*, *B. breve* or *L. lactis* was done using the Qiaprep spin plasmid miniprep kit (Qiagen GmBH, Hilden, Germany). For *B. breve* or *L. lactis* an initial lysis step was incorporated into the plasmid isolation procedure by resuspending cells in lysis buffer supplemented with lysozyme (30 mg ml^{-1}) followed by incubation at 37°C for 30 minutes. Procedures for DNA manipulations were performed essentially as described previously (Sambrook *et al.*, 1989). Restriction enzymes and T4 DNA ligase were used according to the supplier's instructions (Roche Diagnostics, Bell Lane, East Sussex, UK). Synthetic single stranded oligonucleotide primers used in this study are detailed in

(Supplementary table S3.1), and were synthesised by Eurofins (Ebersberg, Germany). Standard PCRs were performed using TaqPCR mastermix (Qiagen GmBH, Hilden, Germany), in a Biometra T3000 thermocycler (Biometra, Goettingen, Germany). PCR products were visualised by EtBR staining following agarose gel electrophoresis (1%). *B. breve* colony PCRs were performed as described previously (O'Connell Motherway *et al.*, 2009). PCR fragments were purified using the Qiagen PCR purification kit (Qiagen GmBH, Hilden, Germany). Electroporation of plasmid DNA into *E. coli* was performed as previously described (Sambrook *et al.*, 1989). Electrotransformation of *B. breve* UCC2003 (Maze *et al.*, 2007) and *L. lactis* (Wells *et al.*, 1993) was performed according to published protocols. The correct orientation and integrity of all plasmid constructs (see also below) were verified by DNA sequencing, performed at Eurofins (Ebersberg, Germany).

Construction of B. breve insertion mutant strains

Internal fragments of *Bbr_1856* (designated here as *melE*) (421 bp, representing codons 242 through to 383 out of the 620 codons of this gene), *Bbr_1857* (designated *melD*) (456 bp, representing codons 92 through to 244 out of the 556 codons of this gene), *Bbr_1860* (designated *melA*) (331 bp, representing codons 230 through to 341 out of the 441 codons of this gene), *Bbr_1867* (designated *rafB*) (394 bp, representing codons 95 through to 226 out of the 429 codons of this gene), *Bbr_1869* (designated *rafA*) (474 bp, representing codons 319 through to 477 out of the 771 codons of this gene) were amplified by PCR using *B. breve* UCC2003 chromosomal DNA as template and the oligonucleotide primer combinations 1856fHd3 and 1856rxba1, 1857fHd3 and 1857rxba1, 1860fHd3 and 1860rxba1, 1867fHd3 and 1867rxba1, and 1869hd3 and 1869Rxba1, respectively. Each of the

generated PCR products was ligated to pORI19, an Ori^+ $RepA^-$ integration plasmid (Law *et al.*, 1995), using HindIII and XbaI restriction sites that were incorporated into the primers for the *melE*, *melD*, *melA*, *rafB* and *rafA* fragment-encompassing amplicons and introduced into *E. coli* EC101 by electroporation. Recombinant *E. coli* EC101 derivatives containing pORI19 constructs were selected on LB agar containing Em, and supplemented with X-gal (5-bromo-4-chloro-3-indolyl-D-galactopyranoside) (40 g ml^{-1}) and 1 mM IPTG.

The expected genetic structure of each of the resulting recombinant plasmids, pORI19-melE (pORI19 containing an internal 421 bp fragment of the *melE* gene), pORI19-melD (pORI19 containing an internal 456 bp fragment of the *melD* gene), pORI19-melA (pORI19 containing an internal 331 bp fragment of the *melA* gene), pORI19-rafB (pORI19 containing an internal 394 bp fragment of the *rafB* gene) and pORI19-rafA (pORI19 containing an internal 474bp fragment of the *rafA* gene), was confirmed by restriction mapping and sequencing prior to subcloning of the Tet resistance antibiotic cassette, *tetW*, from pAM5 (Alvarez-Martin *et al.*, 2007) as a SacI fragment into the unique SacI site in each of the pORI19 derivatives. The orientation of the tetracycline resistance gene in each of the resulting plasmids, pORI19-tet-melE, pORI19-tet-melD, pORI19-tet-melA, pORI19-tet-rafB and pORI19-tet-rafA (naming is consistent with the names of their predecessor plasmids, see above, to which the '-tet-' designation was added), was determined by restriction analysis. The plasmids were subsequently introduced into *E. coli* EC101 pNZ-MBbrI-MBbrII in order to achieve methylation, and transformants were selected based on Cm and Tet resistance. Methylation of the plasmid complement of the obtained transformants in EC101 pNZ-MBbrI-MBbrII was confirmed by their observed insensitivity to PstI restriction (O'Connell Motherway *et al.*, 2009).

Plasmid preparations of methylated pORI19-tet-melE, pORI19-tet-melD, pORI19-tet-melA, pORI19-tet-rafB and pORI19-tet-rafA were then introduced into *B. breve* UCC2003 by electroporation with subsequent selection for transformants on RCA plates supplemented with Tet. Insertion mutants resulting from site-specific homologous recombination were initially confirmed by colony PCR targeting the tetracycline resistance gene *tetW*, followed by a second confirmatory PCR adopting a *tetW*-based primer, either forward or reverse depending on the orientation of *tetW*, in combination with a primer specific for each targeted gene to confirm integration at the correct chromosomal position. In this case, a product would only be obtained if the correctly positioned gene disruption had occurred.

Analysis of global gene expression using B. breve DNA microarrays

Global gene transcription patterns were determined by microarray analysis during growth of *B. breve* UCC2003 on raffinose, stachyose, melibiose, melezitose or sucrose, and these transcriptomes were compared to that obtained from cells that had been grown on ribose as the sole carbohydrate source. All biological replicates were hybridised using a dye-swap strategy. DNA-microarrays containing oligonucleotide primers representing each of the 1864 annotated genes on the genome of *B. breve* UCC2003 were designed by and obtained from Agilent Technologies (Palo Alto, Ca., USA). Methods for cell disruption, RNA isolation, RNA quality control, complementary DNA synthesis and labelling were performed as described previously (Zomer *et al.*, 2009). Labelled cDNA was hybridised using the Agilent Gene Expression hybridisation kit (part number 5188-5242) as described in the Agilent Two-Color Microarray-Based Gene Expression Analysis v4.0 manual (G4140-90050). Following hybridisation, microarrays were washed in accordance with Agilent standard procedures and scanned using an Agilent DNA microarray

scanner (model G2565A). Generated scans were converted to data files with Agilent's Feature Extraction software (Version 9.5). DNA-microarray data were processed as previously described (Garcia De La Nava *et al.*, 2003; van Hijum *et al.*, 2003; van Hijum *et al.*, 2005). Differential expression tests were performed with the Cyber-T implementation of a variant of the *t*-test (Long *et al.*, 2001). A gene was considered differentially expressed when $p < 0.001$, and an expression ratio of >3 or <0.33 relative to the control. The microarray data have been deposited in NCBI's Gene Expression Omnibus and are accessible through GEO Series accession number GSE47448.

Expression and purification of RafA, MelE and MelD

DNA fragments containing the complete (predicted) α-galactosidase-encoding genes, *rafA* and *melE*, or α-glucosidase-encoding gene, *melD*, were generated by PCR amplification from chromosomal DNA of *B. breve* UCC2003 using *Taq* DNA polymerase and primer combinations 1869EcorVF and 1869Xba1R, 1856EcorVF and 1856Xba1R, or 1857Nco1F and 1857Xba1R, respectively (Table S1). NcoI or EcoRV, and XbaI restriction sites were incorporated at the 5' ends of each forward and reverse primer combination, respectively (Table S1). In addition, an in frame His_{10}-encoding sequence was incorporated into each of the forward primers to facilitate downstream protein purification using the Ni-NTA affinity system (Qiagen GmBH, Hilden, Germany). The three generated amplicons were digested with NcoI/EcoRV and XbaI, and ligated into NcoI/ScaI and XbaI-digested nisin-inducible translational fusion plasmids pNZ8048 (which contains an NcoI site) or pNZ8150 (which contains a ScaI site), depending on the restriction sites of the amplified fragment (Mierau *et al.*, 2005). The ligation mixtures were introduced into *L. lactis* NZ9000 (Table 3.1) by electrotransformation, and transformants were selected based

on chloramphenicol resistance. The plasmid content of a number of Cm^r transformants were screened by restriction analysis and the integrity of positively identified clones, namely pNZMelE-His containing gene *melE*, pNZMelD-His containing gene *melD*, and pNZRafA-His containing gene *rafA*, were verified by sequencing.

In order to achieve protein (over)expression and purification, 400 ml of M17 broth supplemented with 0.5 % w/v glucose was inoculated with a 2 % v/v inoculum of a particular *L. lactis* strain, followed by incubation at 30^{o}C until an OD_{600} of 0.5 was reached, at which point protein expression was induced by the addition of purified nisin (5 ng ml^{-1}) followed by continued incubation at 30^{o}C for 90 minutes. Cells were harvested by centrifugation, washed and concentrated 40-fold. Protein purification was performed using a PrepEase® kit specialised for his-tagged protein purification (USB,Germany) Elution fractions were analysed by SDS polyacrylamide gel electrophoresis, as previously described (Laemmli, 1970) on a 12.5 % v/v polyacrylamide gel. After electrophoresis, the gels were fixed and stained with Commassie Brilliant blue to identify fractions containing the purified protein. Rainbow prestained low molecular weight protein markers (New England Biolabs, Herdfordshire, UK) were used to estimate the molecular weight of the purified proteins.

Biochemical characterisation of MelD, MelE and RafA

Determination of the putative α-glucosidase activity of MelD, and the presumed α-galactosidase activity of RafA and MelE were performed essentially as described previously (Börnke *et al.*, 2001). A 50 µl volume of each purified protein (concentration of 0.5 mg ml^{-1}) was incubated with 20 mM MOPS (pH 7.0). In the case of RafA and MelE, 0.1 mg ml^{-1} w/v of α-(1→4)-galactobiose, α-(1→3)-

galactobiose, raffinose, stachyose, melibiose, melezitose, turanose or sucrose was added to 20 mM MOPS buffer as the enzyme substrate in a final volume of 1 ml at 37^{o}C. To analyse the catalytic activities of MelD, 0.1 mg ml^{-1} w/v of α-(1→4)-galactobiose, α-(1→3)-galactobiose, melezitose, maltulose, palatinose, turanose, leucrose, sucrose, raffinose, stachyose or melibiose was added to 20 mM MOPS buffer at pH 7.0 as the enzyme substrate in a final volume of 1 ml and incubated at 37^{o}C. Following incubation, 200 µl samples were taken at 24 hours intervals. Samples were filtered by membrane filtration, using Spin-Xpin-X centrifuge tube filters® (pore size, 0.45 µm; Costar, Corning Inc. NY), and stored at -20°C prior to High-Performance Anion Exchange Chromatography with Pulsed Amperometric Detection (HPAEC-PAD) analysis, see below.

Kinetic constants for MelD using sucrose or turanose, and for RafA using melibiose, raffinose or stachyose were determined by measuring the hydrolysis rates at various substrate concentrations, ranging from 2.5 to 100 mM. Reactions were initiated by the addition of 50µl of purified protein (concentration of 0.5 mg ml^{-1}) in 20 mM MOPS buffer at the optimum pH and temperature determined for each protein, and the reactions were stopped at different time points (up to 6 minutes) by heat treatment at 100°C for 15 minutes. All experiments were performed in duplicate and the amount of released glucose from each disaccharide substrate, namely melibiose, sucrose and turanose, was measured using the glucose hexokinase assay kit (Sigma-Aldrich, Steinheim, Germany) according to the manufacturer's instructions. In the case of raffinose and stachyose hydrolysis, a sucrose assay kit was utilised, in which case enzyme activity is based on the amount of released sucrose as measured according to manufacturer's instructions (Sigma-Aldrich, Steinheim, Germany). Sucrose/turanose or melibiose (50 mM) were used as substrates for the determination

of the pH (pH range of 2.5 to 9.5 was tested) and temperature (a range from 4°C to 60°C was used) optimum. Reactions were initiated by the addition of 50 µl of purified protein (concentration of 0.5 mg ml^{-1}) in 20 mM MOPS buffer at pH 7.0.

High performance anion exchange chromatography with pulsed amperometric detection (hpaec-pad) analysis

For HPAEC-PAD analysis, a Dionex ICS-3000 system (Dionex, Sunnyvale, CA) was used. Carbohydrate fractions (25 µl aliquots) were separated on a CarboPac PA1 (Dionex) analytical-exchange column with dimensions 250 mm x 4 mm with a CarboPac PA1 guard column (Dionex) with dimensions 50 mm x 4 mm and a pulsed electrochemical detector (Dionex ED40) in the pulsed amperometric detection (PAD) mode. Elution was performed at a constant flow-rate of 1.0 ml min^{-1} at 30°C using the following eluents for the analysis (A) 200 mM NaOH, (B) 100 mM NaOH plus 550 mM NaAcetate, and (C) Milli-Q water. The following linear gradient of sodium acetate was used: 100 mM NaOH: 0-50 min, 0 mM; 50-51 min, 16 mM; 51-56 min, 100 mM; 56-61 min, 0 mM. Chromatographic profiles of standard carbohydrates were used for comparison of the results on their breakdown by proteins MelD, MelE and RafA. The Chromeleon software-Version 6.70 (Dionex Corporation) was used for the integration and evaluation of the chromatograms obtained. A 10 mg ml^{-1} stock solution of each of the carbohydrates to be used as reference standards was prepared by dissolving the particular sugar in deionised milli-Q water. The stock solution was sterilised by membrane filtration, using ministart filters (pore size, 0.45 µm; Sartorius AG, Göttingen, Germany), and stored at 4°C.

3.4 RESULTS & DISCUSSION

Carbohydrate-dependent growth analysis of various B. breve strains

Knowledge of carbohydrate metabolism by individual bifidobacterial species or strains is important in order to evaluate the prebiotic potential of particular carbohydrates from a dietary perspective. In order to determine if different strains of *B. breve* possess the ability to utilise the α-glycosidic bond-containing sugars raffinose, stachyose, melibiose and melezitose as sole carbohydrate sources (all of approximately 99 % purity), growth profiles were determined for eleven *B. breve* strains. All tested *B. breve* strains exhibited good growth (final OD_{600} > 1.0) in mMRS supplemented with raffinose, stachyose or melibiose (Fig. 3.1). Six strains exhibited good growth in mMRS supplemented with melezitose, while five strains, namely *B. breve* JCM 7017, *B. breve* 461, *B. breve* 689, *B. breve* NCFB 2257 , *B. longum* subsp. *infantis* ATCC 15697 and *B. breve* NCFB 2258 were not capable of appreciable growth (final OD_{600} < 0.4) on melezitose as a sole carbohydrate source (Fig. 3.1).

Transcriptome analysis of B. breve UCC2003 when grown on raffinose, stachyose or melibiose

In order to identify genes involved in the metabolism of raffinose-related carbohydrates, we investigated differences in global gene expression upon growth of *B. breve* UCC2003 on raffinose, stachyose or melibiose, as compared with growth on ribose (the metabolic pathway for this carbohydrate has previously been characterised in *B. breve* UCC2003, and growth on this sugar has been shown to provide a suitable transcriptomic reference) (Pokusaeva *et al.*, 2010), employing DNA microarray analysis (see materials and Methods). The resulting transcriptomic

data showed that transcription of the contiguous genes *Bbr_1869*, *Bbr_1867*, *Bbr_1866* and *Bbr_1865*, designated here as *rafA, rafB, rafC* and *rafD*, respectively, were significantly upregulated (fold change > 4.0, P < 0.001) in *B. breve* UCC2003 cultures grown on raffinose, stachyose or melibiose, relative to cultures grown on the control carbohydrate ribose, thus implicating this gene cluster in raffinose-related sugar metabolism in *B. breve* UCC2003 (Table 3.2). Furthermore, transcription of four additional genes was upregulated when *B. breve* UCC2003 was grown on raffinose (and on stachyose, although this transcriptional increase was just below the cut-off value of four-fold), but not on melibiose (relative to the ribose transcriptome), namely *Bbr_0026, Bbr_0027, Bbr_0030* and *Bbr_100*, which are predicted to encode two ABC-type permeases, a hypothetical protein, and a putative sucrose phosphorylase, respectively. Since these four genes also exhibit increased transcription when *B. breve* UCC2003 is grown on sucrose (relative to growth on ribose, Table 3.2), we hypothesise that they are involved in the metabolism of this disaccharide, which is released upon removal of the α-galactose moieties from the non-reducing end of raffinose and stachyose, in contrast to melibiose hydrolysis, which results in galactose and glucose release.

Our data on the *raf* cluster of *B. breve* UCC2003 and, by inference, related species suggests that many bifidobacteria metabolise raffinose, and the related sugars stachyose and melibiose, by means of a somewhat different metabolic route compared to that known for *E. coli* (Aslanidis & Schmitt, 1990). Raffinose-type sugar uptake in *B. breve* UCC2003 and other bifidobacteria apparently occurs via an ABC-type transporter system, which is a common way for bifidobacteria to internalise carbohydrates (Maze *et al.*, 2007; O'Connell Motherway *et al.*, 2008; O'Connell Motherway *et al.*, 2011a; Pokusaeva *et al.*, 2009; Pokusaeva *et al.*, 2010;

Pokusaeva *et al.*, 2011b; Ryan *et al.*, 2005). Removal of the α-galactose moiety from raffinose and stachyose results in the release of sucrose, which in the case of *E. coli* is further metabolised via a sucrose hydrolase (Aslanidis & Schmitt, 1990), while in *B. breve* UCC2003 sucrose utilisation appears to occur by a sucrose phosphorylase, as a gene that is predicted to encode such an activity exhibits increased transcription when UCC2003 is grown on sucrose or sucrose-containing sugars (Table 3.2). Sucrose metabolism by sucrose phosphorylase has been characterised in other bifidobacteria, for example *Bifidobacterium adolescentis* DSM20083 (van den Broek *et al.*, 2004), *B. longum* (Kim *et al.*, 2003; Kwon *et al.*, 2007) and *Bifidobacterium animalis* subsp. *lactis* (Trindade *et al.*, 2003). Based on the ability of all examined *B. breve* strains to utilise raffinose-type oligosaccharides (Fig. 3.1), combined with other reports in the literature regarding raffinose utilisation by bifidobacteria and corroborated by the presented comparative genome analysis (see below, Fig. 3.2), raffinose-type sugar utilisation appears to be an ubiquitous property of bifidobacteria, supporting previous publications on the prebiotic potential of (some of) these sugars (Amaretti *et al.*, 2006; Dinoto *et al.*, 2006; Rada *et al.*, 2002; Turroni *et al.*, 2012b).

Although transcription of the *raf* gene cluster of *B. breve* is induced in the presence of either raffinose, melibiose or stachyose, it has been shown that the raffinose system in *Streptococcus pneumoniae* is upregulated in the presence of raffinose, but not melibiose, suggesting differential regulatory mechanisms (Rosenow *et al.*, 1999).

Transcriptome analysis of B. breve UCC2003 when grown on melezitose

In order to identify genes involved in melezitose utilisation by *B. breve* UCC2003, transcriptome analysis was performed upon growth of *B. breve* UCC2003 on this

trisaccharide, and compared with that obtained when grown on ribose. Analysis of transcriptome data obtained from two independent biological replicates revealed that transcription of the contiguous genes *Bbr_1860* through to *Bbr_1856*, which are designated here as *melA, melB, melC, melD* and *melE*, respectively (Fig. 3.2), and which are located in close proximity of the *raf* gene cluster, were significantly upregulated (fold change > 4.0, P < 0.001) in *B. breve* UCC2003 cultures grown on melezitose as compared to cultures grown on ribose (Table 3.2).

These results implicate the *melABCDE* gene cluster in melezitose metabolism in *B. breve* UCC2003. Interestingly, based on previously published comparative genome hybridisation data (O'Connell Motherway *et al.*, 2011b), we noticed that *B. breve* strains that did not exhibit growth on melezitose lack the *melABCDE* gene cluster. Furthermore, and reminiscent of the results obtained for transcriptome analysis when strain UCC2003 was grown on stachyose and raffinose, transcription of four additional genes was upregulated when *B. breve* UCC2003 was grown on melezitose, *Bbr_0026, Bbr_0027, Bbr_0030* and *Bbr_0100* (Table 3.2). As hydrolysis of melezitose produces sucrose if the α-(1→3)-linked glucose is released, the transcriptional induction of these genes is believed to be correlated to sucrose metabolism.

Genetic organisation of the raf and mel gene clusters, and comparison to other available bifidobacterial genomes

Our presumption, based on microarray results, was that the genes of the *raf* gene cluster, schematically depicted in Figure 3.2, are involved in the metabolism of sugars that contain one or more α-(1→6)-linked galactose moieties (e.g. raffinose, stachyose and melibiose). The *rafA* gene, which specifies a putative α-galactosidase

of GH family 27, is a clear homologue of the *B. breve* CECT 7263 α-galactosidase-encoding gene, with which it shares 99 % sequence identity (Jimenez *et al.*, 2012). *B. breve* CECT 7263 also contains genes with high sequence similarity (99-100 % identity) to other genes in the *B. breve* UCC2003 *raf* gene cluster (Fig. 3.2). The *rafA* gene is presumed to be responsible for the breakdown of raffinose, stachyose and melibiose via hydrolysis of the α-(1→6)-glycosidic bond that is common to these carbohydrates, thereby releasing galactose. The presumptive *B. breve* UCC2003 α-galactosidase-encoding gene, *rafA*, is located immediately adjacent to a gene, *Bbr_1868*, designated here as *rafR* and predicted to encode a ROK-type transcriptional regulator. Members of the ROK protein family include both transcriptional repressors and sugar kinases (Titgemeyer *et al.*, 1994), where ROK kinases possess a conserved N-terminal ATP-binding motif, while ROK repressors contain an N-terminal region that includes a canonical helix-turn-helix motif associated with DNA binding (Conejo *et al.*, 2010). BLAST analysis and pfam searches revealed the presence of such a helix-turn-helix motif, thus leading to the prediction that the *rafR* gene product functions as a transcriptional regulator, which may, as based on its genomic location, be involved in regulating raffinose (sugar family)-dependent transcription of the other genes of the *raf* gene cluster. Close homologs of *rafA* and *rafR*, as well as the neighbouring genes *rafB, rafC* and *rafD,* which are predicted to specify a solute binding protein and two permeases, respectively, of a putative ABC-type sugar uptake system, are present in all other publicly available bifidobacterial genomes with the exception of *Bifidobacterium asteroides* PRL2011 and *Bifidobacterium bifidum* PRL2010 (Bottacini *et al.*, 2012; Turroni *et al.*, 2012b) (Fig. 3.2, and data not shown). Interestingly, the *raf* gene cluster does not contain a gene with a predicted ATP-binding protein, which is

required for energy provision of the ABC-type transport system (Schneider & Hunke, 1998). It is presumed that this activity is encoded by an unconnected gene, whose product can function as an ATP-binding component for multiple ABC transporters, which is similar to what has been suggested for other bifidobacterial ABC-type carbohydrate transport systems (O'Connell Motherway *et al.*, 2011a; Pokusaeva *et al.*, 2011b).

The *melABCDE* gene cluster, whose transcription is specifically induced by growth on melezitose, is predicted to specify a solute binding protein (*melA*) and two permease proteins (*melB* and *melC*) (again lacking a gene predicted to encode an ATP-binding protein, representing a similar genetic configuration as that observed for the *raf* gene cluster, see above), an α-glucosidase (*melD*) belonging to GH family 13, and an α-galactosidase/raffinose synthase (*melE*) of GH family 36 (Table 3.2 and Fig. 3.2). The DNA region between the *raf* and *mel* gene clusters contains four additional open reading frames: *Bbr_1863* and *Bbr_1864*, which encode putative LacI-type transcriptional regulators, designated here as $lacI_2$ and $lacI_1$, respectively, and *Bbr_1861* and *Bbr_1862*, which specify a hypothetical protein and a predicted solute binding protein, respectively. The *mel* and *raf* genetic loci including its intervening region genes share high sequence similarity (99-100 %) at the amino acid level and a conserved gene organisation with corresponding regions in the genomes of *B. longum* subsp. *longum* KACC 91563 and *B. longum* subsp. *longum* ATCC 55813. Interestingly, homologs of the *mel* gene cluster (including *Bbr_1861-Bbr_1864*) are lacking in many bifidobacterial genomes, although such genomes do contain clear homologs of the *raf* gene cluster and the *agl1* gene, which in the UCC2003 genome flank the *mel* cluster on either end (Fig. 3.2), suggesting that these bifidobacterial strains lack the (genetic) ability to metabolise melezitose.

Prevalence and genetic organisation of raf and mel gene clusters in other bacteria

The genetic organisation of the bifidobacterial *raf* locus and the raffinose metabolic pathway appears to be quite different from those of various other bacterial species. In *E. coli* (Aslanidis *et al.*, 1989), *Klebsiella pneumoniae* (Hama & Wilson, 1992), *Enterobacter cloacae* (Okazaki *et al.*, 1997) and *Citrobacter freundii* (Shimamoto *et al.*, 2001), among others, raffinose uptake is specified by a single permease-encoding gene, which is co-transcribed with two additional genes, that specify an α-galactosidase and a sucrose hydrolase. As mentioned above, uptake of raffinose and related sugars in bifidobacteria appears to be performed by a dedicated ABC-type transport system, the corresponding genes of which are in close proximity of, although they are not co-transcribed with the α-galactosidase-specifying gene, due to their opposing genetic orientation (Fig. 3.2). Furthermore, the metabolic product of this α-galactosidase activity, i.e. sucrose, is in bifidobacteria apparently not metabolised by a sucrose hydrolase, but by a sucrose phosphorylase, encoded by a gene that is not genetically linked to the *raf* locus.

Raffinose metabolism has also been investigated in *Lactobacillus plantarum* ATCC 8014, in which the putative raffinose permease-encoding gene is immediately followed by the α-galactosidase-specifying gene (designated *melA*), while preceded by *galM*, a putative microbial galactose-1-epimerase (Silvestroni *et al.*, 2002). Interestingly, two genes are located downstream of *melA*, though on the opposite strand that encode the two subunits of a heterodimeric β-galactosidase, which shows that the genetic organisation and content of the raffinose utilisation cluster of *Lb. plantarum* differs substantially from that found in bifidobacteria. Recently it has been shown by transcriptome analyses that *B. animalis* subsp. *lactis* Bl-04 possesses a raffinose utilisation cluster including three predicted α-glycosidases, exhibiting a

transcription profile that is consistent with our observations for *B. breve* (Andersen *et al.*, 2013).

Very little information is available regarding melezitose metabolism in other microorganisms. Melezitose metabolism has been investigated in the yeast *Saccharomyces cerevisiae* (Hwang & Lindegren, 1964), where a melezitose metabolising system was identified, including an α-glucosidase or melezitase that was shown to be capable of hydrolysing palatinose, turanose, maltose, sucrose and melezitose.

Construction and phenotypes of mutants carrying individual disruptions in the rafA, rafB, melA, melD and melE genes

In order to establish if disruption of particular genes of the *raf* gene cluster of *B. breve* UCC2003 affects the resulting strain's ability to metabolise raffinose, stachyose and/or melibiose, insertion mutants were made in *rafB*, which is predicted to encode a solute binding protein implicated in the internalisation of the above mentioned α-galactose-containing sugars, and *rafA*, the presumed α-(1→6)-galactosidase-encoding gene. The resulting strains were designated UCC2003-RafB and UCC2003-RafA, respectively (see Materials and Methods). *B. breve* UCC2003-RafA and *B. breve* UCC2003-RafB were analysed for their ability to grow on mMRS supplemented with either raffinose, stachyose or melibiose as the sole carbon source. As expected and in contrast to the wild type, the *B. breve* insertion mutants exhibited an inability to grow on raffinose, stachyose or melibiose as the sole carbon source. All strains retained their ability to utilise glucose as a sole carbon source (Fig. 3.3A). These results not only show that the *rafA* gene is required for growth on raffinose, but also that a mutation in *rafB* causes this growth-deficient phenotype, implying that the *rafBCD* genes encode an ABC-type transport system responsible

for the internalisation of α-galactose-containing oligosaccharides. Our results corroborate findings of a recent study, in which a deletion in a homologue of the *rafA* gene of *Bifidobacterium longum* 105-A (designated *agl*) was shown to cause loss of α-galactosidic activity and a growth deficiency on raffinose or melibiose (Hirayama *et al.*, 2012).

In order to establish if disruption of particular genes from the *mel* cluster results in the loss of *B. breve* UCC2003's ability to metabolise melezitose, insertion mutants were made in *melE*, which is predicted to encode an α-galactosidase, resulting in strain *B. breve* UCC2003-MelE, in *melD*, which is predicted to encode an α-glucosidase, resulting in strain *B. breve* UCC2003-MelD and in *melA*, which encodes a predicted solute binding protein, resulting in strain *B. breve* UCC2003-MelA (see Materials and Methods). These mutants, i.e. *B. breve* UCC2003-MelE, UCC2003-MelD and UCC2003-MelA, were then analysed for their ability to grow on mMRS supplemented with melezitose as the sole carbon source. As expected and in contrast to the wild type, growth of *B. breve* strain UCC2003-MelA on melezitose was severely reduced compared to that of the wild type strain, indicating that this gene is indeed required for melezitose catabolism. Mutant strain *B. breve* UCC2003-MelD was also impaired in growth on melezitose as compared to the wild type strain UCC2003, though to a lesser extent than the UCC2003-MelA strain, and this growth defect is particularly obvious when comparing the corresponding growth profiles of these three strains on melezitose (Fig. 3.3B). The less severe growth deficiency (as compared to the MelA strain) may be due to the presence of other α-glucosidases produced by *B. breve* UCC2003 that partially compensate for the *melD* mutation. Interestingly, mutant strain UCC2003-MelE was not affected in its ability to grow on melezitose, indicating that MelE is not required for growth on this sugar (data not

shown). Wild type UCC2003 and the three *mel* mutants did not exhibit any differences in their ability to utilise glucose as a sole carbon source.

Purification, characterisation and substrate specificity of recombinantly produced MelD, MelE and RafA

In order to analyse the glycosyl hydrolase functions of MelD, MelE and RafA, we purified and biochemically characterised these three predicted sugar-degrading enzymes. All of the overproduced proteins purified well and in soluble form, and MelE, MelD and RafA exhibited molecular masses of approximately 68.2 kDa , 62.4 kDa and 84.1 kDa (inclusive of the His_6-tag), respectively, when analysed by SDS-PAGE (Results not shown).

Purified recombinant MelD protein was shown to fully hydrolyse melezitose into glucose and fructose (Fig. 3.4A I), thereby demonstrating that this protein has both α-(1→2) and α-(1→3) glucosyl hydrolase activities. This was further confirmed by our findings that both sucrose [α-D-Glc*p*-(1→2)-β-D-Fru*f*] and turanose [α-D-Glc*p*-(1→3)-β-D-Fru*f*] are hydrolysed by MelD to glucose and fructose (Fig. 3.4 II and III, respectively). In contrast, MelD was unable, at least under the conditions tested, to hydrolyse leucrose [α-D-Glc*p*-(1→5)-β-D-Fru*f*], palatinose [α-D-Glc*p*-(1→6)-β-D-Fru*f*] or maltulose [α-D-Glc*p*-(1→4)-β-D-Fru*f*] (results not shown). These results therefore show that MelD is an α-glucosidase that exhibits clearly different substrate specificities compared to the two previously characterised α-glucosidases encoded by *B. breve* UCC2003, Agl1 and Agl2, which do not possess the ability to hydrolyse leucrose, melezitose or sucrose, but possess such hydrolytic activity against turanose, maltulose and palatinose (Pokusaeva *et al.*, 2009). The preferred substrates for Agl1 and Agl2 are thus carbohydrates that contain either an α-(1→6) or an α-(1→1)

glucosidic bond (Pokusaeva *et al.*, 2009), while MelD is an α-glucosidase with hydrolytic activity against α-(1→2) and α-(1→3) glucosidic bonds, which makes the latter enzyme different from previously characterised bifidobacterial α-glucosidases (Degnan & Macfarlane, 1994; Igaue, 1983; van den Broek *et al.*, 2003).

The putative α-galactosidase MelE failed to exhibit hydrolytic activity towards raffinose, stachyose or melibiose (results not shown), at least under the conditions tested. However, when two synthetic disaccharides, namely α-(1→4) and α-(1→3) galactobiose were assayed, MelE was shown to be capable of hydrolysing both disaccharides (Supplementary figure S3.1). The finding that the *melE* gene encodes an α-galactosidase and is present and co-transcribed in a melezitose-induced gene cluster suggests that this locus is also involved in the metabolism of (a) melezitose-related carbohydrate(s), which contain(s) one or more alpha-galactose moieties linked through α-(1→4) and/or α-(1→3) glycosidic bonds. This sugar, like melezitose, may also be present in honeydew, which is secreted by aphids while feeding on the sugar-rich phloem of their host plants, to acquire the amino acids they need for growth and reproduction (Wool *et al.*, 2006). Such honeydew oligosaccharides, which contain large amounts of melezitose and erlose, have been shown to possess prebiotic potential as they increase bifidobacterial and lactobacilli numbers in an *in vitro* fermentation system (Sanz *et al.*, 2005).

Purified RafA was shown to hydrolyse stachyose, raffinose and melibiose (Fig. 3.4B I, II and III), to produce sucrose and galactose, and cleaving melibiose to its monosaccharide constituents glucose and galactose, thereby confirming that RafA functions as an α-galactosidase. RafA was also tested for its ability to hydrolyse melezitose, sucrose, and a range of sucrose isomers, but did not exhibit hydrolytic activity against any of these carbohydrates (results not shown). However, it did show

hydrolytic activity against synthetic α-(1→4) and α-(1→3) galactobiose, showing that RafA exhibits a rather broad substrate specificity (Supplementary table S3.2). Previously, an α-galactosidase from *B. bifidum* JCM 1254 was shown to be capable of hydrolysing α-(1→3)-linked galactose in a branched blood group B antigen trisaccharide, however, this α-galactosidase did not possess the ability to hydrolyse α-(1→4)-galactosidic linkages (Wakinaka *et al.*, 2013). Analysis of currently sequenced bifidobacterial genomes shows that most of the bifidobacterial strains encode at least one α-galactosidase-encoding gene (Pokusaeva *et al.*, 2011a). Various α-galactosidases, capable of catalysing hydrolysis of various α-galacto-oligosaccharides, have been studied in four bifidobacterial species to date (Garro *et al.*, 1994; Goulas *et al.*, 2009; Leder *et al.*, 1999; Van Laere *et al.*, 1999; Wakinaka *et al.*, 2013; Zhao *et al.*, 2008). For example, in the case of *B. adolescentis*, it was shown that the α-galactosidase hydrolyses α-(1→6)-galactosidic bonds from raffinose and stachyose, but also the α-(1→4) and α-(1→3) bonds of two galactobiose substrates (Leder *et al.*, 1999; Van Laere *et al.*, 1999), which is consistent with our findings for *B. breve* UCC2003. Interestingly, the α-galactosidase of *B. breve* 203 was shown to have the ability to synthesise a trisaccharide (Gal-α-(1→4)-Gal-α-(1→6)-Glc) using melibiose as a substrate (Zhao *et al.*, 2008). Furthermore, the α-galactosidase encoded by *B. bifidum* JCM 1254 was observed to be capable of hydrolysing α-(1→3)-linked galactose in a branched blood group B antigen trisaccharide, although this enzyme cannot hydrolyse α-(1→4)-galactosidic linkages (Wakinaka *et al.*, 2013).

The obtained MelD, MelE and RafA-mediated carbohydrate hydrolysis results and presumed degradation pathways have been summarised in supplementary Figure S3.2.

Determination of kinetic details of MelD and RafA

In order to determine the kinetic parameters of the MelD and RafA enzymes, we characterised these two glycosyl hydrolases using the substrates against which they had shown hydrolytic activity (with the exception of α-(1→4) and α-(1→3) galactobiose, for which we did not have sufficient amounts to perform such studies). When sucrose and turanose were used as substrates, the optimum temperature and pH values for MelD activity were determined to be 30°C and 7.5, respectively, while in the case of RafA in combination with either of the substrates raffinose, stachyose or melibiose the optima were 42°C and 6.0, respectively.

Kinetic studies were performed to determine V_{max} and K_m values, as well as the rate constants (k_{cat}) and catalytic efficiencies (k_{cat}/K_m), for MelD using sucrose and turanose as substrates, and for RafA employing raffinose, stachyose and melibiose as substrates (Table 3.3). MelD, as shown above, exhibits hydrolytic activity against both α-(1→3) and α-(1→2) glucosidic linkages. In order to investigate if MelD exhibited any hydrolytic preference for either of these linkages, we looked at the ability of MelD to hydrolyse sucrose and turanose. The obtained data indicate that the preferred bond cleaved by MelD is the α-(1→2) linkage present in sucrose, as MelD hydrolyses this bond with a higher efficacy compared to the α-(1→3) linkage present in turanose (Table 3.3). In order to investigate the preferred substrate of RafA we determined the kinetic parameters of this enzyme related to its hydrolytic activity towards melibiose, raffinose and stachyose. The preferred substrate was melibiose, followed by raffinose. However, it is important to note that stachyose contains two α-(1→6) linkages with one of these present between two galactose moieties, while the other being present between galactose and glucose. The hydrolysis of stachyose by RafA is likely to lead to the generation of galactose and

raffinose, the latter again representing a substrate for RafA, as demonstrated by us and previously reported (Garro *et al.*, 1999). In our kinetic experiments we assessed stachyose hydrolysis by measuring the release of sucrose, which would be generated directly by hydrolysis of stachyose to sucrose and a galacto-disaccharide and indirectly by initial hydrolysis of stachyose to galactose and raffinose, followed by hydrolysis of the raffinose to sucrose and galactose. The presence of multiple substrates means that the use of standard kinetic techniques to measure stachyose hydrolysis is not absolutely correct. However, the kinetic values calculated provide a reasonably accurate representation of the substrate preference of the enzyme.

A previously published kinetic study of the *E. coli* K12 α-galactosidase (Schmid & Schmitt, 1976) revealed that this enzyme at an optimal pH 7.2 exhibits Km values of 3.2 mM for melibiose and 60 mM for raffinose. Since we observed Km values of 8.15 for melibiose and 14.55 for raffinose at an optimal pH of 6.0 for the *B. breve* RafA, it is clear that individual α-galactosidases, despite having the same substrate specificities, may still exhibit varying kinetic properties.

3.5 CONCLUDING REMARKS

Bifidobacteria are believed to play an important role in the fermentation of non-digestible carbohydrates in the lower gastrointestinal tract, and consistent with this notion is the prediction that a sizable proportion of the average bifidobacterial genome is dedicated to carbohydrate metabolism (Ventura *et al.*, 2007a; Ventura *et al.*, 2007b). Over 50 bifidobacterial carbohydrases have been studied to date, and various carbohydrate utilisation pathways have been characterised in *B. breve* UCC2003, such as those dedicated to the metabolism of fructose, galactan, starch, ribose, palatinose, cellodextrin, and fructo-oligosaccharides (Leahy *et al.*, 2005; Maze *et al.*, 2007; Mierau *et al.*, 2005; O'Connell Motherway *et al.*, 2011b; O'Riordan & Fitzgerald, 1999; Oishi *et al.*, 2008; Pokusaeva *et al.*, 2011b).

The data assembled in this study provide significant information on the ability of various *B. breve* strains to grow on a number of plant-derived α-glucose and α-galactose-containing oligosaccharides, while also identifying the corresponding genes involved in the metabolism of such sugars. Two adjacent genetic loci dedicated to the utilisation of raffinose-containing carbohydrates and melezitose in *B. breve* UCC2003 were identified, encoding a novel α-glucosidase, specified by *melD*, and two α-galactosidases, specified by *rafD* and *melE*, as well as presumed ABC-type uptake systems for their corresponding carbohydrate substrates.

Bifidobacteria appear to encode both common carbohydrate utilisation pathways (e.g. for the metabolism of raffinose-family sugars), as well as strain/species-specific pathways (e.g. melezitose metabolism). Such pathways may be a reflection of common elements in the diet of the host of such bacteria, while it may also allow certain species/strains the capacity to effectively colonise the gut or obtain higher

numbers when the host's diet contains more specialised carbohydrates. In a recent study (van Zanten *et al.*, 2012) *B. animalis* subsp. *lactis* Bl-04 was shown to be selectively increased in a model system of the human colon 10 to 100-fold on melibiose, xylobiose, raffinose and maltotriose, indicating that these carbohydrates have the potential to serve as a prebiotic.

It is possible that the ability to lose or acquire a carbohydrate utilisation system is similar to that of *Lb. plantarum*, an organism which appears to acquire, shuffle, substitute or delete carbohydrate utilisation systems in response to niche requirements, and which makes it a "natural metabolic engineer" (Siezen & van Hylckama Vlieg, 2011). Further investigations will be required to determine how and to what extent specific carbohydrate utilisation abilities facilitate certain bifidobacterial species and strains to colonise and persist in the gastrointestinal tract of their host, and the importance of host diet in this regard.

3.6 ACKNOWLEDGEMENTS

This publication has emanated from research conducted with the financial support of Science Foundation Ireland (SFI) under Grant Numbers 07/CE/B1368, 08/SRC/B13404 and SFI/12/RC/2273. KJOC was supported by a postgraduate fellowship funded through the Tomar trust, while MOCM is a recipient of a HRB postdoctoral fellowship (Grant no. PDTM/20011/9). We also sincerely thank J. Thompson (National Institutes of Health, Bethesda, MD) and N. Emphadinhas (Universidade de Coimbra, Coimbra, Portugal) for supplying various sucrose isomers, Francesca Bottacini for her bioinformatics input, and Dr Stephen Cunningham, Dr. Marian Keane and Prof. Lokesh Joshi of the Alimentary Glycoscience Research Cluster, NUIG, Ireland, for kindly supplying α-(1→4) and α-(1→3) galactobiose.

3.7 TABLES AND FIGURES

Table 3.1 Bacterial strains and plasmids used in this study

Strain/ plasmid	Relevant characteristics	Reference or source
E.coli strains		
EC101	Cloning host, repA+ kmr	(Law *et al.*, 1995)
EC101-pNZ-M.Bbrll+Bbr111	EC101 harbouring pNZ8048 derivative containing bbrllM and bbrlllM	(Mary O Connell-Motherway *et al.*, 2009)
L.lactis strains		
NZ9000	MG1363,nisin-inducible overexpression host; pepN::nisRK	(de Ruyter *et al.*, 1996)
NZ9700	Nisin producing strain	(de Ruyter *et al.*, 1996)
B. breve strains		
UCC2003	Isolate from a nursling stool	(Maze *et al.*, 2007)
NCFB 2257	Isolate from infant intestine	NCFB
NCTC 11815	Isolate from infant intestine	NCTC
NCFB 2258	Isolate from infant intestine	NCFB
NCIMB 8815	Isolate from infant faeces	NCIMB
JCM 7017	Isolate from human faeces	JCM
JCM 7019	Isolate from infant faeces	JCM
UCC2005	Isolate from nursling stool	(Pokusaeva *et al.*, 2009)
Yakult	Isolate from nursling stool	(Oishi *et al.*, 2008)
Nizo 658	Isolate from nursling stool	Nizo
461	Isolate from infant/adult faeces	PRL
689	Isolate from infant/adult faeces	PRL
12L	Mothers milk	PRL
UCC2003-MelE	pORI19-tet-1856 insertion mutant of UCC2003	This study

Strain/ plasmid	Relevant characteristics	Reference or source
UCC2003-MelD	pORI19-tet-1857 insertion mutant of UCC2003	This study
UCC2003-RafB	pORI19-tet-1867 insertion mutant of UCC2003	This study
UCC2003-RafA	pORI19-tet-1869 insertion mutant of UCC2003	This study
UCC2003-MelA	pORI19-tet-1860 insertion mutant of UCC2003	This study
Plasmids	pORI19-tet-1857 insertion mutant of UCC2003	This study
pORI19	Emr , repA- , ori+ , cloning vector	(Law *et al.*, 1995)
pORI19-tet-MelE	Internal 421bp fragment of melE and tetW cloned in pORI19	This study
pORI19-tet-MelD	Internal 456bp fragment of melD and tetW cloned in pORI19	This study
pORI19-tet-MelA	Internal 331bp fragment of MelA and tetW cloned in pORI19	This study
pORI19-tet-RafA	Internal 474bp fragment of RafA and tetW cloned in pORI19	This study
pORI19-tet-RafB	Internal 394bp fragment of RafB and tetW cloned in pORI19	This study
pAM5	pBC1-puC19-Tcr	(Alvarez-Martin *et al.*, 2007)
pNZ8048	Cm^r; nisin-inducible translational fusion vector	(de Ruyter *et al.*, 1996)
pNZ8150	Cm^r; nisin-inducible translational fusion vector	(Mierau *et al.*, 2005)
pNZMelE-His	MelE with his tag cloned downstream of nisin inducible promoter on pNZ8048	This study
pNZMelD-His	MelD with his tag cloned downstream of nisin inducible promoter on pNZ8048	This study
pNZRafA-His	RafA with his tag cloned downstream of nisin inducible promoter on pNZ8150	This study

JCM: Japan Collection of Microorganisms; NIZO: Nizo Food Research; NCFB: National Collection of Food Bacteria; NCIMB: National Collection of Industrial and Marine Bacteria; NCTC: National Collection of Type Cultures; PRL: Culture collection of probiogenomics University of Parma.

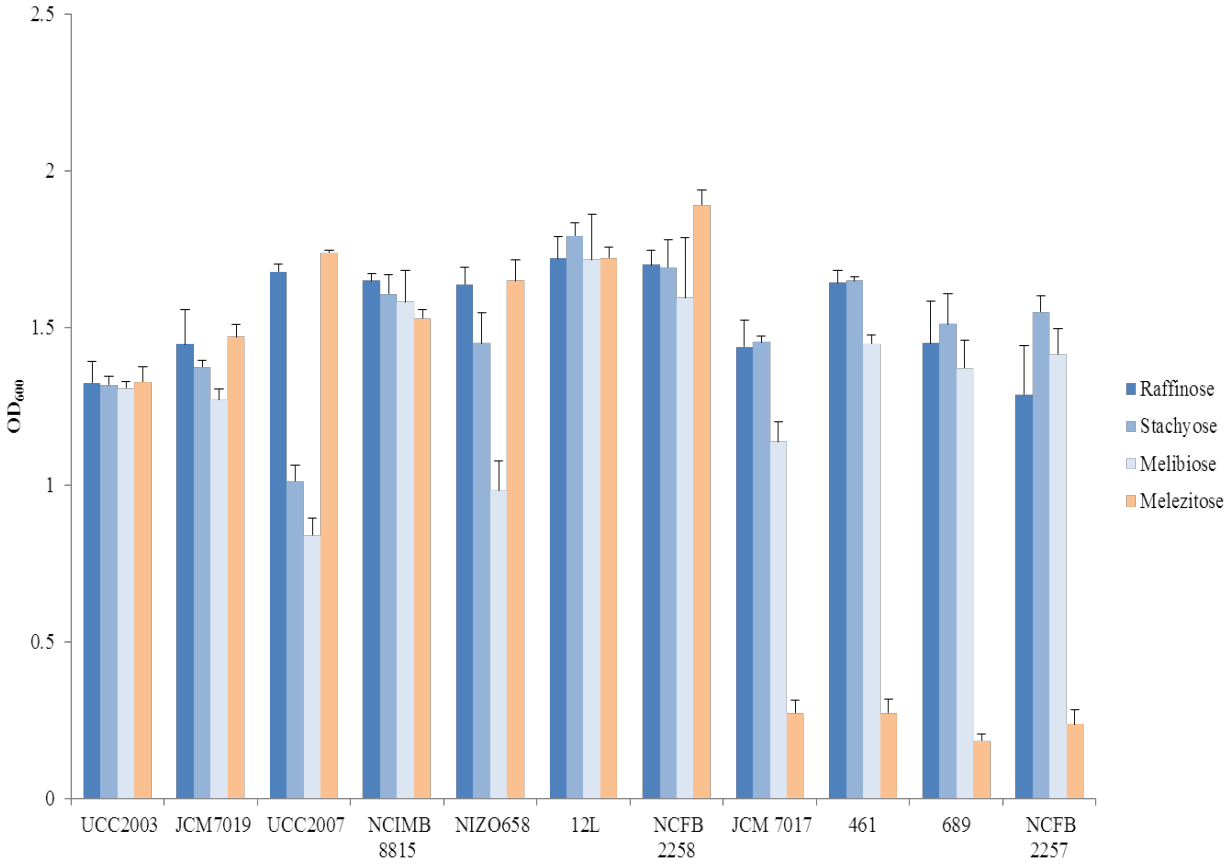

Figure 3.1 Final OD600 values following 16 hours of growth of various wild type *B. breve* strains grown on 1 % raffinose, 1 % stachyose, 1 % melibiose and 1 % melezitose. The results are mean values obtained from three separate experiments.

Table 3.2 Carbohydrate-dependent transcriptional upregulation of specific genes as based on transcriptome comparative analysis using *B. breve* UCC2003 grown on 1% raffinose, stachyose and melibiose as compared to growth on ribose. Microarray data were obtained using *B. breve* UCC2003 grown on 1% raffinose, stachyose and melibiose, and using array data obtained when grown on ribose as control.

			Fold upregulation during growth on:				
Gene ID.	Gene Name	Function	Stachyose	Raffinose	Melibiose	Melezitose	Sucrose
Bbr_0026		Permease protein of ABC transporter system for sugars	_	18.55	_	4.66	23.73
Bbr_0027		Permease protein of ABC transporter system for sugars	_	13.98	_	4.55	44.95
Bbr_0030		Conserved hypothetical protein	28.53	145.75	_	86.54	216.5
Bbr_0100	*SPase*	Sucrose phosphorylase	74.26	74.26	_	128.18	41.71
Bbr_1855	*agl1*	Alpha glucosidase	_	20.43	10.74	31.95	_
Bbr_1856	*melE*	Raffinose synthase or seed inhibition protein	_	_	_	72.76	_
Bbr_1857	*melD*	Alpha glucosidase	_	_	_	94.33	_
Bbr_1858	*melC*	Permease protein of ABC transporter system for sugars	_	_	_	850.82	_
Bbr_1859	*melB*	Permease protein of ABC transporter system for sugars	_	_	_	679.07	_
Bbr_1860	*melA*	Solute binding protein of ABC transporter system for sugars	_	_	_	724.89	_
Bbr_1865	*rafD*	Raffinose transport system permease protein	21.66	98.36	105.07	_	_
Bbr_1866	*rafC*	Raffinose transport system permease protein	49.86	175.51	153.8	_	_
Bbr_1867	*rafB*	Raffinose-binding protein	24.39	86.16	92.5	_	_
Bbr_1868	*rafR*	Transcriptional repressor, ROK family	_	_	_	_	_
Bbr_1869	*rafA*	Alpha-galactosidase	9.23	10.53	8.07	_	_

Fold cut off point is 4 fold, p value < 0.001, and values below cut-off are indicated in the Table as dashes.

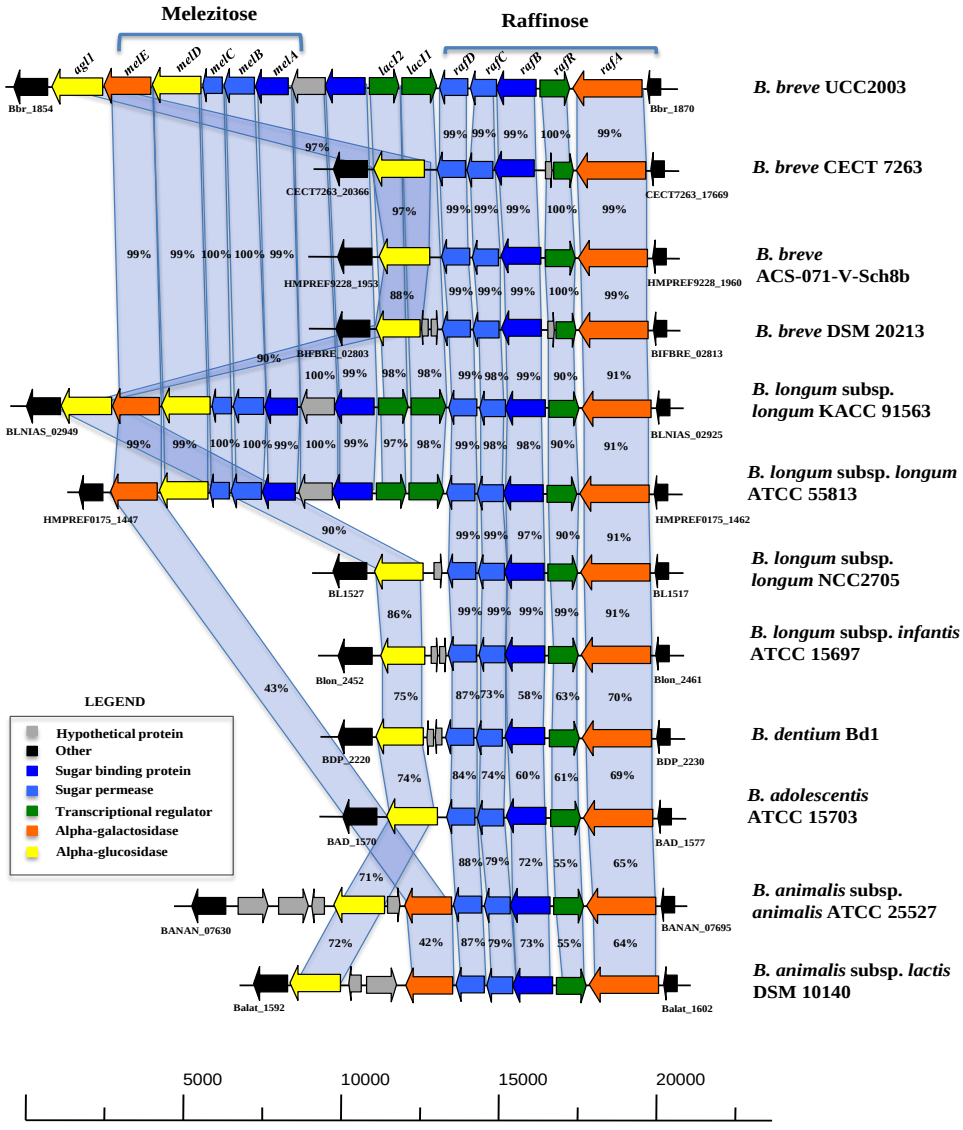

Figure 3.2 Comparison of the melezitose, raffinose and stachyose gene clusters of *B. breve* UCC2003 with corresponding putative melezitose, raffinose and stachyose utilisation loci from other bifidobacteria. Each solid arrow indicates an open reading frame (ORF). The lengths of the arrows are proportional to the length of the predicted ORF and the gene locus name, which is indicative of its putative function, is indicated above the arrow. Orthologs are marked with the same colour while the amino acid identity of each predicted protein is indicated as a percentage relative to its equivalent protein encoded by *B. breve* UCC2003.

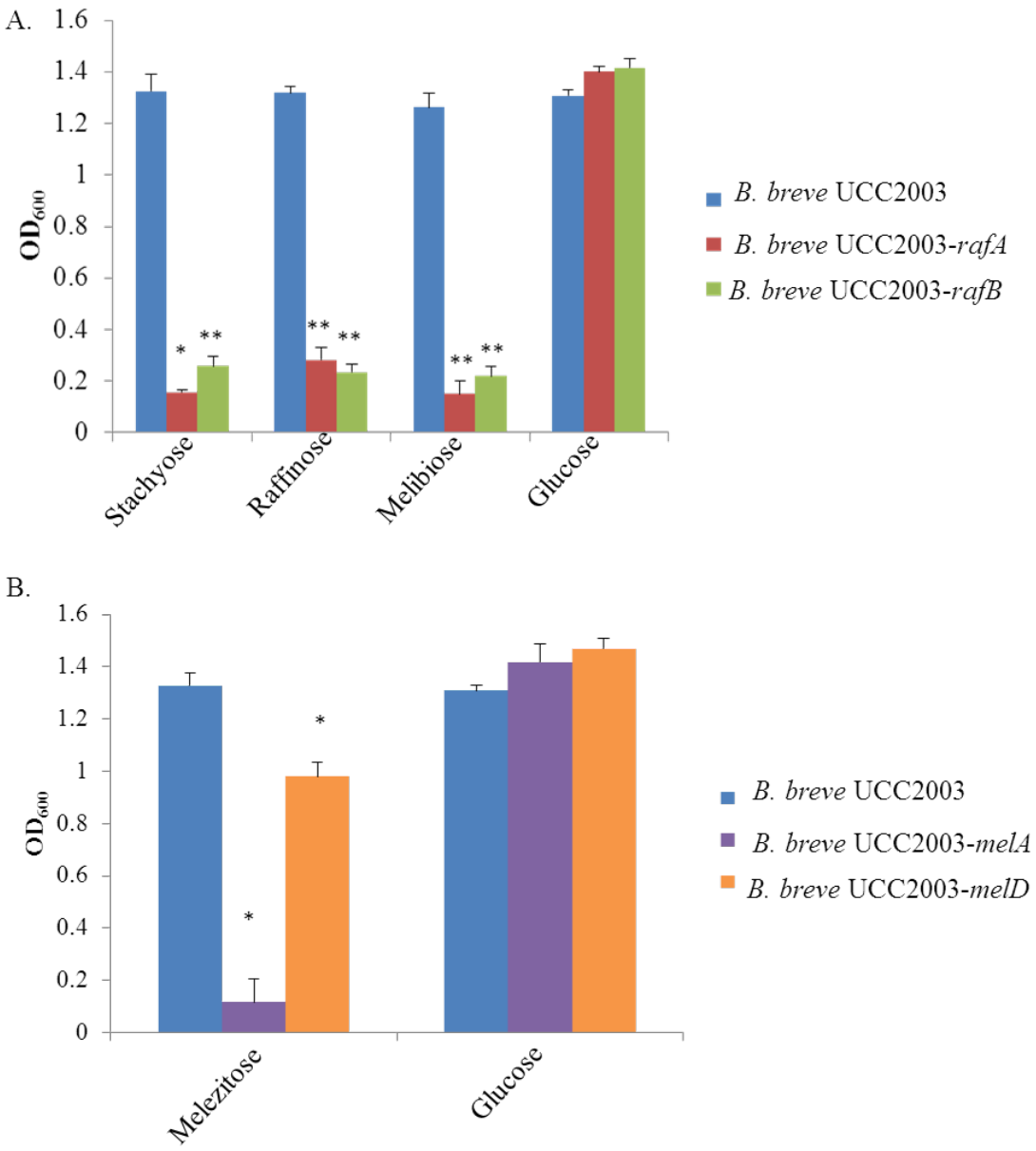

Figure 3.3 Panel A; Final OD600 after 16 hours UCC2003 and insertion mutants UCC2003-RafB (raffinose binding protein) and UCC2003-RafA (α-galactosidase) on 1 % stachyose, raffinose, melibiose and glucose. The results are mean values obtained from three separate experiments. *P< 0.01, **P< 0.001. **Panel B.** Final OD600 following 16 hours of growth of *B. breve* UCC2003, *B. breve* UCC2003-MelA (solute binding protein) and *B. breve* UCC2003-MelD (α-(1→3)-glucosidase) on 1 % melezitose and glucose. The results are mean values obtained from three experiments,*P<0.01.

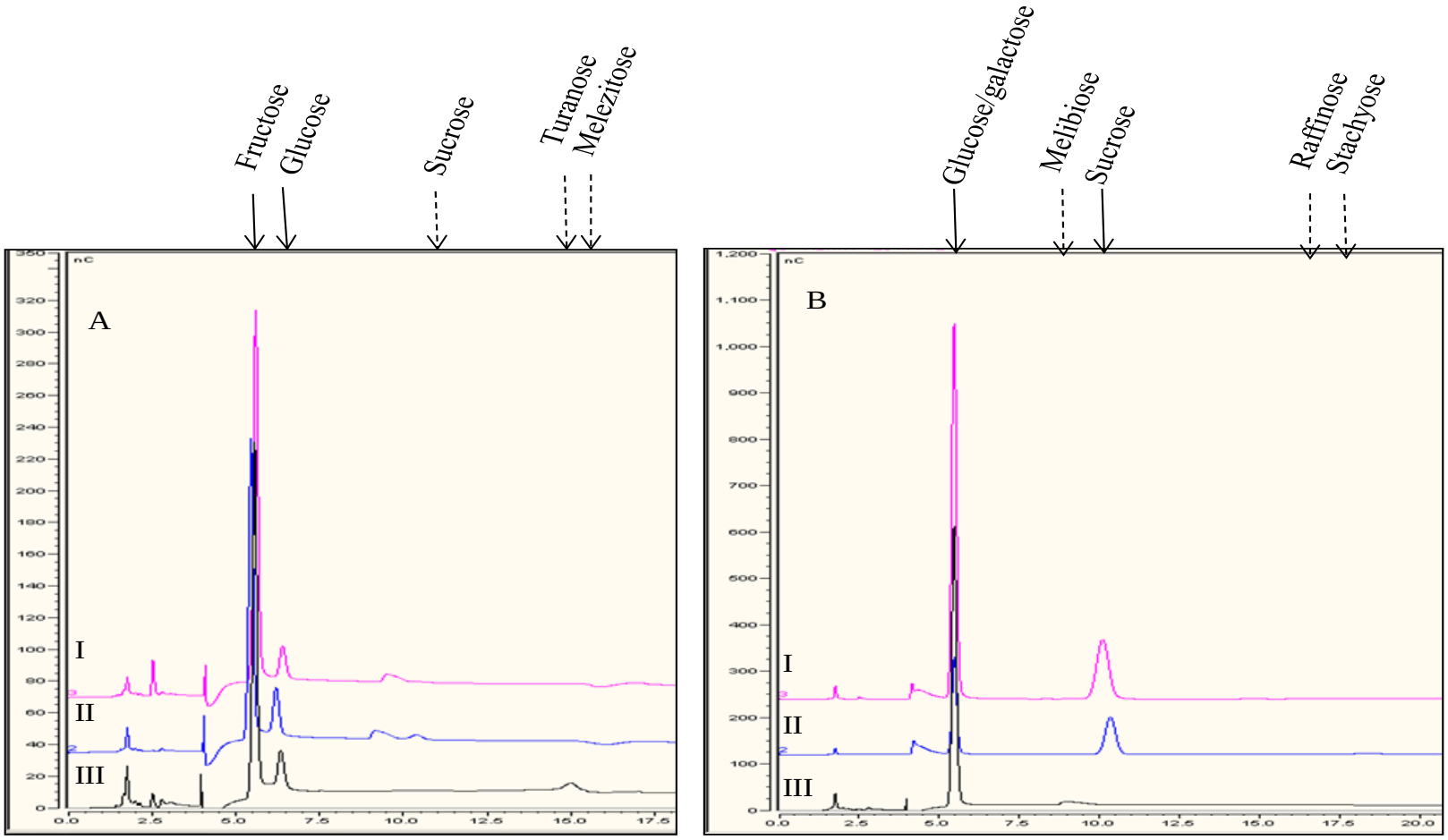
Fructose
Glucose
Sucrose
Turanose
Melezitose
A
I
II
III
Glucose/galactose
Melibiose
Sucrose
Raffinose
Stachyose
B
I
II
III

Figure 3.4 Panel A: HPAEC-PAD analysis indicating melezitose and turanose breakdown 0.1 mg ml^{-1} by the purified recombinant protein MelD in 20 mM MOPS buffer (pH 7.0) over 24 hours. Graph I: Melezitose incubated with MelD, where liberation of glucose and fructose is visible as chromatographic peaks eluted at 6.25 and 5.5 minutes, respectively; Graph II: Sucrose incubated with MelD, where liberation of glucose and fructose is visible as chromatographic peaks eluted at 6.25 and 5.5 minutes, respectively; graph III: Turanose incubated with MelD where liberation of glucose and fructose is visible as chromatographic peaks eluted at 6.25 and 5.5 minutes, respectively; Breakdown products indicated with black compound arrows. Chromatographic positions of carbohydrate standards are indicated by dashed arrows above the chromatogram. **Panel B:** HPAEC-PAD analysis indicating stachyose, raffinose and melibiose breakdown by the purified recombinant protein RafA in 20 mM MOPS buffer (ph. 7.0) over 24 hours. Graph I: Stachyose incubated with RafA, where liberation of galactose and sucrose is visible as chromatographic peaks eluted at 5.75 and 10.5 minutes, respectively; Graph II: Raffinose incubated with RafA, where release of galactose and sucrose is detected as chromatographic peaks eluted at 5.75 and 10.5, minutes respectively; Graph III: Melibiose incubated with RafA, where hydrolysis of this substrate to glucose and galactose is visible as a single chromatographic peak eluted at 5.75 minutes. Breakdown products indicated with black compound arrows. Chromatographic positions of carbohydrate standards are indicated by dashed arrows above the chromatogram.

Table 3.3 Kinetic study of MelD and RafA, V_{max} and K_m values for the substrate sucrose including the rate constant (k_{cat}) and catalytic efficiency (k_{cat}/K_m).

Substrate	Protein	Vmax (μmol $min^{-1}mg^{-1}$)	Km(Mm)	Kcat(s^{-1})	kcat/km($Mm^{-1}s^{-1}$)
Melibiose	RafA	384.305±0.42	8.15±1.7	542.79±0.63	66.6±0.07
Raffinose	RafA	187.5±2.1	14.55±1.9	264.815±2.9	18.2±2.6
Stachyose	RafA	11.09±.55	3.045±.51	16.665±0.784	5.47±1.14
Sucrose	MelD	7.75±0.07	19.3±0.28	8.0725±0.07	0.4175±0.007
Turanose	MelD	12.66±1.59	91.93±1.07	13.16±1.61	0.14±0.01

All values are means from two experiments ± standard errors.

3.8 SUPPLEMENTARY DATA

Table S3.1 Oligonucleotide primers used in this study

Primer	Sequence [a]	Size
1856fHd3	TGCGGA**AAGCTT**TGTACTTGGGATTCATTG	421bp
1856rxba1	CTATGC**TCTAGA**CGAAATCGACGCCGGCT	
1857fHd3	TGCGGA**AAGCTT**GAACGCGGTATCGGCATCATC	457bp
1857rxba1	CTATGC**TCTAGA**GCACGGCGTCGCGGTCCCAGA	
1867fHd3	TGCGGA**AAGCTT**GCCGGACGTAATCACCTTC	394bp
1867rxba1	CTATGC**TCTAGA**CGACTCGGGAACCAGCGTAC	
1869Fhd3	TGCGGA**AAGCTT**CTGTTCGGCTCCTTCGGCG	474bp
1869Rxba1	CTATGC**TCTAGA**GTCAGGTCGAGCACTTGC	
1857-confirm	CATTCGAGCTGCCACACA	2.6bp
Tetwsal1F	TCAGCT**GTCGAC**ATGCTCATGTACGGTAAG	
1857-confirm	ACTTTCACCAACAATGACTG	2.6kb
Tetwsal1F	TCAGCT**GTCGAC**ATGCTCATGTACGGTAAG	
1867-confirm	CAACGCCTAGACTGAAGC	2.6kb
Tetwsal1F	TCAGCT**GTCGAC**ATGCTCATGTACGGTAAG	
1869-confirm	GCGTTTCCGGTGCACTCGATG	2.6kb
Tetwsal1F	TCAGCT**GTCGAC**ATGCTCATGTACGGTAAG	
Tetwsal1F	TCAGCT**GTCGAC**ATGCTCATGTACGGTAAG	2.7kb
Tetwsal1R	GCGACG**GTCGAC**CATTACCTTCTGAAACATA GACAAG**GATATC**ATGCATCACCATCACCATCACCATCACCA	
1856EcorVF	TCACATTGCCGAAACCTGTCC	1.9bp
1856Xba1	GACAAG**TCTAGA**AACTGCGTTGGTGTGCCGG TGCACG**CCATGG**CTCATCACCATCACCATCACCATCACCAT	
1857Nco1F	CACACTTTCACCAACAATGACTG	1.6kb
1857Xba1R	CTATGC**TCTAGA**TTATGCTGGCAAGGCCCAGACG GACAAG**GATATC**ATGCATCACCATCACCATCACCATCACCA	
1869EcorVF	TCACGCCCAGAACACCGCCACC	2.3kb
1869Xba1	CTATGC**TCTAGA**TCAGCTTAACGCCACGGCCTTG	

[a] Sequences of restriction enzyme sites are indicated in bold

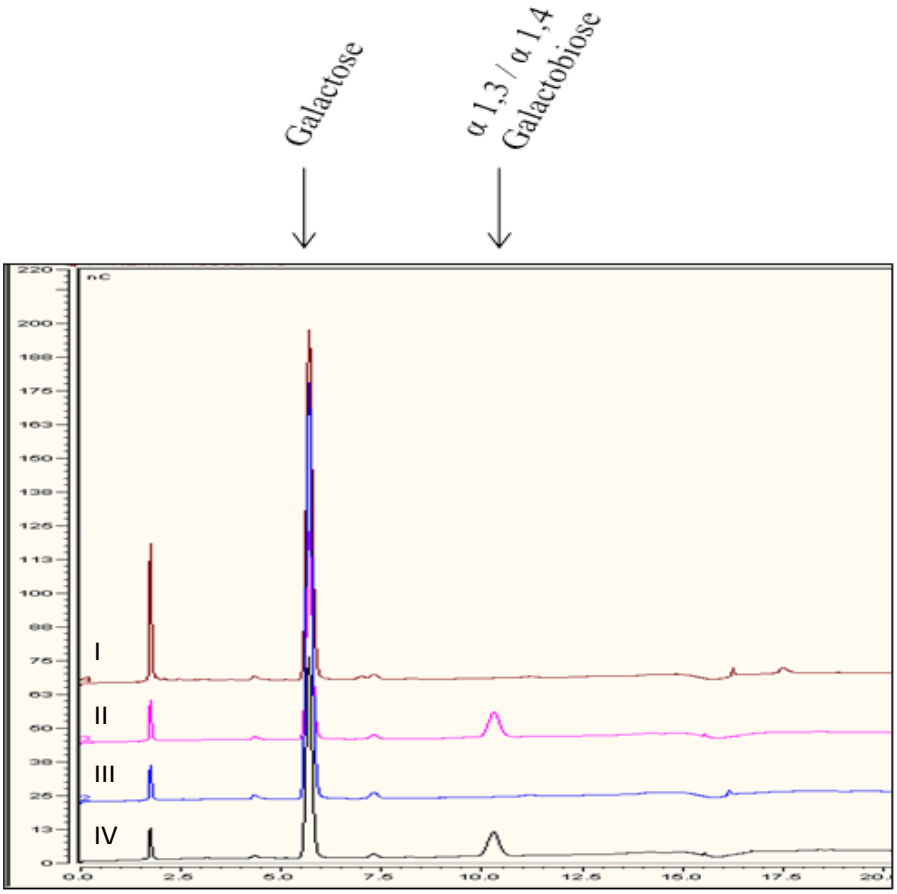

Figure S3.1 HPAEC-PAD analysis indicating α 1,3 and α 1,4 galactobiose breakdown 0.1 mg ml^{-1} by the purified recombinant proteins MelE and RafA in 20 mM MOPS buffer (pH 7.0) over 24 hours. Graph I and Graph II: α-(1→3) and α-(1→4) galactobiose incubated with MelE, where liberation of galactose is visible as a chromatographic peak eluted at 5.9 minutes; Graph III and graph IV: α-(1→3) and α-(1→4) galactobiose incubated with MelE, where liberation of galactose is visible as a chromatographic peak eluted at 5.9 minutes; Breakdown products indicated with black compound arrows. Chromatographic positions of carbohydrate standards are indicated by dashed arrows above the chromatogram.

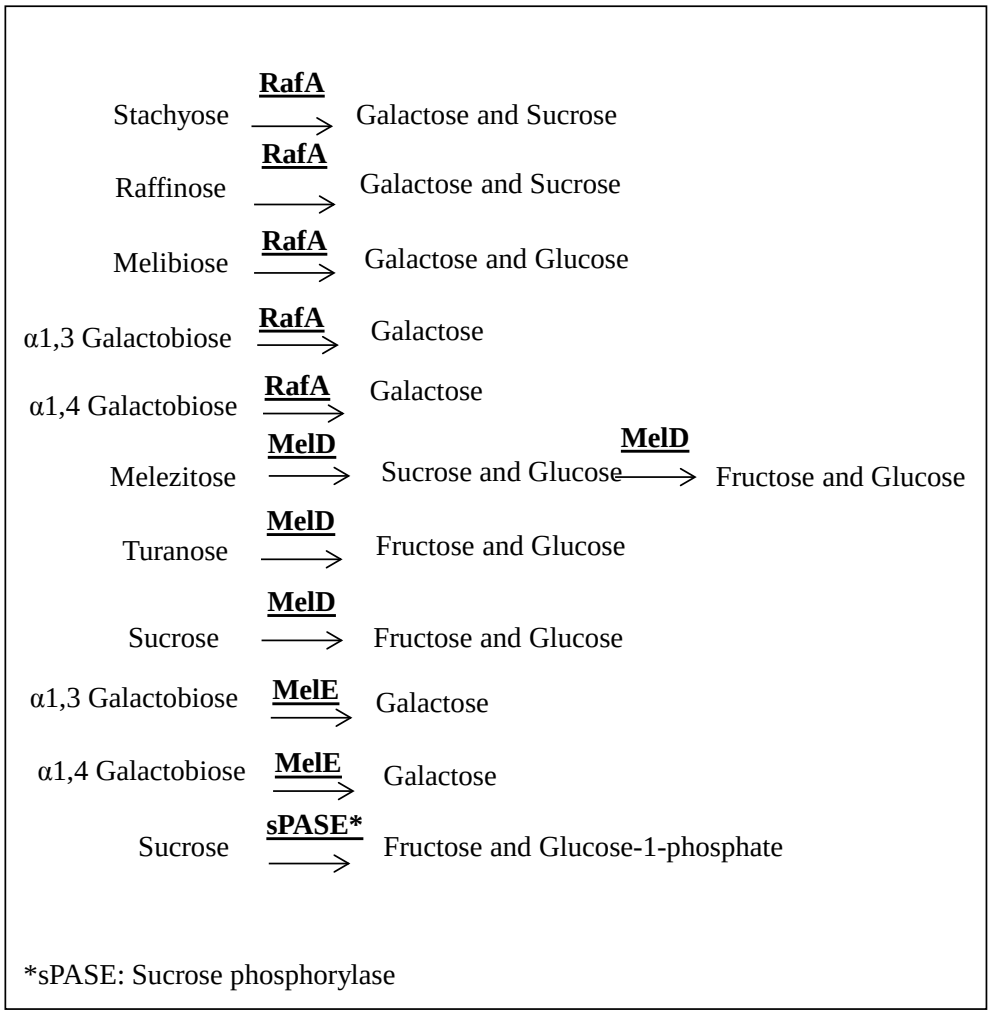

Figure S3.2 Proposed sugar metabolism pathways

3.9 REFERENCES

Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. (1990). Basic local alignment search tool. *Journal of Molecular Biology* **215**, 403-410.

Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J., Zhang, Z., Miller, W. & Lipman, D. J. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Research* **25**, 3389-3402.

Alvarez-Martin, P., O'Connell-Motherway, M., van Sinderen, D. & Mayo, B. (2007). Functional analysis of the pBC1 replicon from *Bifidobacterium catenulatum* L48. *Applied Microbiology and Biotechnology* **76**, 1395-1402.

Amaretti, A., Tamburini, E., Bernardi, T., Pompei, A., Zanoni, S., Vaccari, G., Matteuzzi, D. & Rossi, M. (2006). Substrate preference of *Bifidobacterium adolescentis* MB 239: compared growth on single and mixed carbohydrates. *Applied Microbiology and Biotechnology* **73**, 654-662.

Andersen, J., Barrangou, R., Hachem, M., Lahtinen, S., Goh, Y., Svensson, B. & Klaenhammer, T. (2013). Transcriptional analysis of oligosaccharide utilisation by *Bifidobacterium lactis* Bl-04. *BMC Genomics* **14**, 312.

Aslanidis, C., Schmid, K. & Schmitt, R. (1989). Nucleotide sequences and operon structure of plasmid-borne genes mediating uptake and utilisation of raffinose in *Escherichia coli*. *Journal of Bacteriology* **171**, 6753-6763.

Aslanidis, C. & Schmitt, R. (1990). Regulatory elements of the raffinose operon: nucleotide sequences of operator and repressor genes. *Journal of Bacteriology* **172**, 2178-2180.

Bacon, J. S. & Dickinson, B. (1957). The origin of melezitose: a biochemical relationship between the lime tree (Tilia spp.) and an aphis (Eucallipterus tiliae L.). *The Biochemical Journal* **66**, 289-297.

Börnke, F., Hajirezaei, M. & Sonnewald, U. (2001). Cloning and Characterisation of the Gene Cluster for Palatinose Metabolism from the Phytopathogenic *Bacterium Erwinia rhapontici. Journal of Bacteriology* **183**, 2425-2430.

Bottacini, F., Milani, C., Turroni, F., Sánchez, B., Foroni, E., Duranti, S., Serafini, F., Viappiani, A., Strati, F., Ferrarini, A., Delledonne, M., Henrissat, B., Coutinho, P., Fitzgerald, G.F., Margolles, A., van Sinderen D. & Ventura, M. (2012). *Bifidobacterium asteroides* PRL2011 genome analysis reveals clues for colonisation of the insect gut. *PLoS One* **7**, e44229.

Conejo, M. S., Thompson, S. M. & Miller, B. G. (2010). Evolutionary bases of carbohydrate recognition and substrate discrimination in the ROK protein family. *Journal of Molecular Evolution* **70**, 545-556.

De Man, J. C., Rogosa, M. & Sharpe, M. E. (1960). A medium for the cultivation of *lactobacilli*. *Journal of Applied Microbiology* **23**, 130-135.

de Ruyter, P. G., Kuipers, O. P., Beerthuyzen, M. M., van Alen-Boerrigter, I. & de Vos, W. M. (1996). Functional analysis of promoters in the nisin gene cluster of *Lactococcus lactis*. *Journal of Bacteriology* **178**, 3434-3439.

Degnan, B. A. & Macfarlane, G. T. (1994). Synthesis and activity of a-glucosidase produced by *Bifidobacterium pseudolongum*. *Current Microbiology* **29**, 43-47.

Dinoto, A., Suksomcheep, A., Ishizuka, S., Kimura, H., Hanada, S., Kamagata, Y., Asano, K., Tomita, F. & Yokota, A. (2006). Modulation of rat cecal microbiota by administration of raffinose and encapsulated *Bifidobacterium breve*. *Applied and Environmental Microbiology* **72**, 784-792.

French, D. (1954). The Raffinose Family of Oligosaccharides. In *Advances in Carbohydrate Chemistry*, pp. 149-184. Edited by L. W. Melville: Academic Press.

Garcia De La Nava, J., Santaella, D. F., Alba, J. C., Carazo, J. M., Trelles, O. & Pascual-Montano, A. (2003). Engene: the processing and exploratory analysis of gene expression data. *Bioinformatics* **19**, 657-658.

Garro, M. S., de Giori, G. S., de Valdez, G. F. & Oliver, G. (1994). α-D-Galactosidase (EC 3.2.1.22) from *Bifidobacterium longum*. *Letters in Applied Microbiology* **19**, 16-19.

Garro, M. S., de Valdez, G. F., Oliver, G. & de Giori, G. S. (1999). Hydrolysis of soya milk oligosaccharides by *Bifidobacterium longum* CRL 849. *European Journal of Nutrition* **208**, 57-59.

Gibson, G. R. & Roberfroid, M. B. (1995). Dietary modulation of the human colonic microbiota: introducing the concept of prebiotics. *The Journal of Nutrition* **125**, 1401-1412.

Goulas, T., Goulas, A., Tzortzis, G. & Gibson, G. R. (2009). A novel alpha-galactosidase from *Bifidobacterium bifidum* with transgalactosylating properties: gene molecular cloning and heterologous expression. *Applied Microbiology and Biotechnology* **82**, 471-477.

Hama, H. & Wilson, T. H. (1992). Primary structure and characteristics of the melibiose carrier of *Klebsiella pneumoniae*. *Journal of Biological Chemistry* **267**, 18371-18376.

Hirayama, Y., Sakanaka, M., Fukuma, H., Murayama, H., Kano, Y., Fukiya, S. & Yokota, A. (2012). Development of a double-crossover markerless gene deletion system in *Bifidobacterium longum*: functional analysis of the alpha-galactosidase gene for raffinose assimilation. *Applied and Environmental Microbiology* **78**, 4984-4994.

Hudson, C. S. (1946). Melezitose and Turanose. In *Advances in Carbohydrate Chemistry*, pp. 1-36. Edited by M. L. W. W.W. Pigman & P. Stanley: Academic Press.

Hwang, D. S. & Lindegren, C. C. (1964). Palatinose Element of the Receptor of the Melezitose Locus in Saccharomyces. *Nature* **203**, 791-792.

Igaue, I. W., H.; Oda, T.; Oyamada, K. (1983). Enzymes of human intestinal bacteria active in degrading maltitol, 2: Some characteristics of maltitol-hydrolysing enzyme (alpha-glucosidase) from a strain of *Bifidobacterium adolescentis*. *Journal of the Agricultural Chemical Society of Japan* **57**, 995-999.

Jimenez, E., Villar-Tajadura, M. A., Marin, M., Fontecha, J., Requena, T., Arroyo, R., Fernandez, L. & Rodriguez, J. M. (2012). Complete genome sequence of *Bifidobacterium breve* CECT 7263, a strain isolated from human milk. *Journal of Bacteriology* **194**, 3762-3763.

Kim, M., Kwon, T., Joo Lee, H., Heon Kim, K., Kyun Chung, D., Eog Ji, G., Byeon, E.-S. & Lee, J.-H. (2003). Cloning and expression of sucrose phosphorylase gene from *Bifidobacterium longum* in *E. coli* and characterisation of the recombinant enzyme. *Biotechnology Letters* **25**, 1211-1217.

Kwon, T., Kim, C. T. & Lee, J.-H. (2007). Transglycosylation of ascorbic acid to ascorbic acid 2-glucoside by a recombinant sucrose phosphorylase from *Bifidobacterium longum*. *Biotechnology Letters* **29**, 611-615.

Laemmli, U. K. (1970). Cleavage of Structural Proteins during the Assembly of the Head of Bacteriophage T4. *Nature* **227**, 680-685.

Law, J., Buist, G., Haandrikman, A., Kok, J., Venema, G. & Leenhouts, K. (1995). A system to generate chromosomal mutations in *Lactococcus lactis* which allows fast analysis of targeted genes. *Journal of Bacteriology* **177**, 7011-7018.

Leahy, S. C., Higgins, D. G., Fitzgerald, G. F. & van Sinderen, D. (2005). Getting better with bifidobacteria. *Journal of Applied Microbiology* **98**, 1303-1315.

Leder, S., Hartmeier, W. & Marx, S. P. (1999). Alpha-galactosidase of *Bifidobacterium adolescentis* DSM 20083. *Current Opinion in Microbiology* **38**, 101-106.

Liu, Z., Jiang, Z., Zhou, K., Li, P., Liu, G. & Zhang, B. (2007). Screening of bifidobacteria with acquired tolerance to human gastrointestinal tract. *Anaerobe* **13**, 215-219.

Long, A. D., Mangalam, H. J., Chan, B. Y., Tolleri, L., Hatfield, G. W. & Baldi, P. (2001). Improved statistical inference from DNA microarray data using analysis of variance and a Bayesian statistical framework. Analysis of global gene expression in *Escherichia coli* K12. *Journal of Biological Chemistry* **276**, 19937-19944.

Macfarlane, G. T., Steed, H. & Macfarlane, S. (2008). Bacterial metabolism and health-related effects of galacto-oligosaccharides and other prebiotics. *Journal of Applied Microbiology* **104**, 305-344.

Maze, A., O'Connell-Motherway, M., Fitzgerald, G. F., Deutscher, J. & van Sinderen, D. (2007). Identification and characterisation of a fructose phosphotransferase system in *Bifidobacterium breve* UCC2003. *Applied and Environmental Microbiology* **73**, 545-553.

Mierau, I., Leij, P., van Swam, I., Blommestein, B., Floris, E., Mond, J. & Smid, E. (2005). Industrial-scale production and purification of a heterologous protein in *Lactococcus lactis* using the nisin-controlled gene expression system NICE: The case of lysostaphin. *Microbial Cell Factories* **4**, 15.

Minami, Y., Yazawa, K., Tamura, Z., Tanaka, T. & Yamamoto, T. (1983). Selectivity of utilisation of galactosyl-oligosaccharides by bifidobacteria. *Chemical & Pharmaceutical Bulletin* **31**, 1688-1691.

Miyake, T., Watanabe, K., Watanabe, T. & Oyaizu, H. (1998). Phylogenetic analysis of the genus *Bifidobacterium* and related genera based on 16S rDNA sequences. *Microbiology and Immunology* **42**, 661-667.

O'Connell Motherway, M., Fitzgerald, G. F., Neirynck, S., Ryan, S., Steidler, L. & van Sinderen, D. (2008). Characterisation of ApuB, an extracellular type II amylopullulanase from *Bifidobacterium breve* UCC2003. *Applied and Environmental Microbiology* **74**, 6271-6279.

O'Connell Motherway, M., O'Driscoll, J., Fitzgerald, G. F. & van Sinderen, D. (2009). Overcoming the restriction barrier to plasmid transformation and targeted mutagenesis in *Bifidobacterium breve* UCC2003. *Microbial Biotechnology* **2**, 321-332.

O'Connell Motherway, M., Fitzgerald, G. F. & van Sinderen, D. (2011a). Metabolism of a plant derived galactose-containing polysaccharide by *Bifidobacterium breve* UCC2003. *Microbial Biotechnology* **4**, 403-416.

O'Connell Motherway, M., Zomer, A., Leahy, S. C., Reunanen, J., Bottacini, F., Claesson, M.J., O'Brien, F., Flynn, K., Casey, P.G., Munoz, J.A., Kearney, B., Houston, A.M., O' Mahony, C., Higgins, DG., Shanahan, F., Palva, A., de Vos, W.M., Fitzgerald, G.F., Ventura, M., O'Toole, P.W. & van Sinderen D. (2011b). Functional genome analysis of *Bifidobacterium breve* UCC2003 reveals type IVb tight adherence (Tad) pili as an essential and conserved host-colonisation factor. *Proceedings of the National Academy of Sciences of the United States of America* **108**, 11217-11222.

O'Riordan, K. & Fitzgerald, G. F. (1999). Molecular characterisation of a 5.75-kb cryptic plasmid from *Bifidobacterium breve* NCFB 2258 and determination of mode of replication. *FEMS Microbiology Letters* **174**, 285-294.

Oishi, K., Sato, T., Yokoi, W., Yoshida, Y., Ito, M. & Sawada, H. (2008). Effect of probiotics, *Bifidobacterium breve* and *Lactobacillus casei*, on bisphenol A exposure in rats. *Bioscience, Biotechnology, and Biochemistry* **72**, 1409-1415.

Okazaki, N., Jue, X. X., Miyake, H., Kuroda, M., Shimamoto, T. & Tsuchiya, T. (1997). A melibiose transporter and an operon containing its gene in *Enterobacter cloacae*. *Journal of Bacteriology* **179**, 4443-4445.

Pokusaeva, K., O'Connell-Motherway, M., Zomer, A., Fitzgerald, G. F. & van Sinderen, D. (2009). Characterisation of two novel alpha-glucosidases from *Bifidobacterium breve* UCC2003. *Applied and Environmental Microbiology* **75**, 1135-1143.

Pokusaeva, K., Neves, A. R., Zomer, A., O'Connell-Motherway, M., MacSharry, J., Curley, P., Fitzgerald, G. F. & van Sinderen, D. (2010). Ribose utilisation by the human commensal *Bifidobacterium breve* UCC2003. *Microbial Biotechnology* **3**, 311-323.

Pokusaeva, K., Fitzgerald, G. & van Sinderen, D. (2011a). Carbohydrate metabolism in Bifidobacteria. *Genes and Nutrition* **6**, 285-306.

Pokusaeva, K., O'Connell-Motherway, M., Zomer, A., Macsharry, J., Fitzgerald, G. F. & van Sinderen, D. (2011b). Cellodextrin utilisation by *Bifidobacterium breve* UCC2003. *Applied and Environmental Microbiology* **77**, 1681-1690.

Rada, V., Bartonova, J. & Vlkova, E. (2002). Specific growth rate of bifidobacteria cultured on different sugars. *Folia microbiologica* **47**, 477-480.

Rehms, H. & Barz, W. (1995). Degradation of stachyose, raffinose, melibiose and sucrose by different tempe-producing *Rhizopus fungi*. *Applied Microbiology and Biotechnology* **44**, 47-52.

Rosenow, C., Maniar, M. & Trias, J. (1999). Regulation of the alpha-galactosidase activity in *Streptococcus pneumoniae*: characterisation of the raffinose utilisation system. *Genome research* **9**, 1189-1197.

Rutherford, K., Parkhill, J., Crook, J., Horsnell, T., Rice, P., Rajandream, M. A. & Barrell, B. (2000). Artemis: sequence visualisation and annotation. *Bioinformatics* **16**, 944-945.

Ryan, S. M., Fitzgerald, G. F. & van Sinderen, D. (2005). Transcriptional Regulation and Characterisation of a Novel β-Fructofuranosidase-Encoding Gene from *Bifidobacterium breve* UCC2003. *Applied and Environmental Microbiology* **71**, 3475-3482.

Sambrook, J., Fritsch, E. F. & Maniatis, T. (1989). *Molecular cloning: a laboratory manual*: Cold Spring Harbour Laboratory.

Sanz, M. L., Polemis, N., Morales, V., Corzo, N., Drakoularakou, A., Gibson, G. R. & Rastall, R. A. (2005). *In Vitro* Investigation into the Potential Prebiotic

Activity of Honey Oligosaccharides. *Journal of Agricultural and Food Chemistry* **53**, 2914-2921.

Scardovi, V. & Trovatelli, I. D. (1965). The fructose-6-phosphate shunt as a peculiar pattern of hexose degradation in the genus *Bifidobacterium*. *Annals of Microbiology* **15**, 19-29.

Schell, M. A., Karmirantzou, M., Snel, B., Vilanova, D., Berger, B., Pessi, G., Zwahlen, M.C., Desiere, F., Bork, P., Delley, M., Pridmore, R.D. & Arigoni, F. (2002). The genome sequence of *Bifidobacterium longum* reflects its adaptation to the human gastrointestinal tract. *Proceedings of the National Academy of Sciences of the United States of America* **99**, 14422-14427.

Schmid, K. & Schmitt, R. (1976). Raffinose Metabolism in *Escherichia coli* K12. *European Journal of Biochemistry* **67**, 95-104.

Schneider, E. & Hunke, S. (1998). ATP-binding-cassette (ABC) transport systems: Functional and structural aspects of the ATP-hydrolysing subunits/domains. *FEMS Microbiology Reviews* **22**, 1-20.

Shimamoto, T., Shimamoto, T., Xu, X. J., Okazaki, N., Kawakami, H. & Tsuchiya, T. (2001). A cryptic melibiose transporter gene possessing a frameshift from *Citrobacter freundii*. *Journal of Biochemistry* **129**, 607-613.

Siezen, R. & van Hylckama Vlieg, J. (2011). Genomic diversity and versatility of *Lactobacillus plantarum*, a natural metabolic engineer. *Microbial Cell Factories* **10**, S3.

Silvestroni, A., Connes, C., Sesma, F., De Giori, G. S. & Piard, J. C. (2002). Characterisation of the melA locus for alpha-galactosidase in *Lactobacillus plantarum. Applied and Environmental Microbiology* **68**, 5464-5471.

Stanton, C., Ross, R. P., Fitzgerald, G. F. & van Sinderen, D. (2005). Fermented functional foods based on probiotics and their biogenic metabolites. *Current Opinion in Biotechnology* **16**, 198-203.

Tannock, G. W. (1997). Probiotic properties of lactic-acid bacteria: plenty of scope for fundamental R & D. *Trends in Biotechnology* **15**, 270-274.

Terzaghi, B. E. & Sandine, W. E. (1975). Improved medium for lactic *Streptococci* and their bacteriophages. *Applied Microbiology* **29**, 807-813.

Tissier, H. (1900). Recherchers sur la flora intestinale normale pathologique du nourisson. *Thesis, University of Paris, Paris France.*

Titgemeyer, F., Reizer, J., Reizer, A. & Saier, M. H. (1994). Evolutionary relationships between sugar kinases and transcriptional repressors in bacteria. *Microbiology* **140**, 2349-2354.

Trindade, M. I., Abratt, V. R. & Reid, S. J. (2003). Induction of Sucrose Utilisation Genes from *Bifidobacterium lactis* by Sucrose and Raffinose. *Applied and Environmental Microbiology* **69**, 24-32.

Turroni, F., Ribbera, A., Foroni, E., van Sinderen, D. & Ventura, M. (2008). Human gut microbiota and bifidobacteria: from composition to functionality. *Antonie van Leeuwenhoek* **94**, 35-50.

Turroni, F., Peano, C., Pass, D. A., Foroni, E., Severgnini, M., Claesson, M. J., Kerr, C., Hourihane, J., Murray, D., Fuligni, F., Gueimonde, M., Margolles, A., De Bellis, G., O'Toole, P. W., van Sinderen, D., Marchesi, J. R., & Ventura, M. (2012a). Diversity of Bifidobacteria within the Infant Gut Microbiota. *PLoS ONE* **7**, e36957.

Turroni, F., Strati, F., Foroni, E., Serafini, F., Duranti, S., van Sinderen, D. & Ventura, M. (2012b). Analysis of Predicted Carbohydrate Transport Systems Encoded by *Bifidobacterium bifidum* PRL2010. *Applied and Environmental Microbiology* **78**, 5002-5012.

van den Broek, L. A., Struijs, K., Verdoes, J. C., Beldman, G. & Voragen, A. G. (2003). Cloning and characterisation of two alpha-glucosidases from *Bifidobacterium adolescentis* DSM20083. *Applied Microbiology and Biotechnology* **61**, 55-60.

van den Broek, L. A., van Boxtel, E. L., Kievit, R. P., Verhoef, R., Beldman, G. & Voragen, A. G. (2004). Physico-chemical and transglycosylation properties of recombinant sucrose phosphorylase from *Bifidobacterium adolescentis* DSM20083. *Applied Microbiology and Biotechnology* **65**, 219-227.

van Hijum, S. A. F. T., Garcia De La Nava, J., Trelles, O., Kok, J. & Kuipers, O. P. (2003). MicroPreP: a cDNA microarray data pre-processing framework. *Applied Bioinformatics* **2**, 241-244.

van Hijum, S. A. F. T., De, J. A., Baerends, R. J., Karsens, HA., Kramer, NE., Larsen, R., den Hengst, C.D., Albers, C.J., Kok, J. & Kuipers, O.P. (2005). A generally applicable validation scheme for the assessment of factors involved in reproducibility and quality of DNA-microarray data. *BMC Genomics* **6**, 77.

Van Laere, K. M., Hartemink, R., Beldman, G., Pitson, S., Dijkema, C., Schols, H. A. & Voragen, A. G. (1999). Transglycosidase activity of *Bifidobacterium adolescentis* DSM 20083 alpha-galactosidase. *Applied Microbiology and Biotechnology* **52**, 681-688.

van Zanten, G. C., Knudsen, A., Roytio, H., Forssten, S., Lawther, M., Blennow, A., Lahtinen, S.J., Jakobsen, M., Svensson, B. & Jespersen, L. (2012). The effect of selected synbiotics on microbial composition and short-chain fatty acid production in a model system of the human colon. *PLoS One* **7**, e47212.

Ventura, M., Canchaya, C., Tauch, A., Chandra, G., Fitzgerald, G. F., Chater, K. F. & van Sinderen, D. (2007a). Genomics of *Actinobacteria*: Tracing the Evolutionary History of an Ancient Phylum. *Microbiology and Molecular Biology Reviews* **71**, 495-548.

Ventura, M., O'Connell-Motherway, M., Leahy, S., Moreno-Munoz, J. A., Fitzgerald, G. F. & van Sinderen, D. (2007b). From bacterial genome to functionality; case bifidobacteria. *International Journal of Food Microbiology* **120**, 2-12.

Ventura, M., Turroni, F., Motherway, M. O., MacSharry, J. & van Sinderen, D. (2012). Host-microbe interactions that facilitate gut colonisation by commensal bifidobacteria. *Trends in Microbiology* **20**, 467-476.

Wakinaka, T., Kiyohara, M., Kurihara, S,. Hirata, A., Chaiwangsri, T., Ohnuma, T., Fukamizo, T., Katayama, T., Ashida, H. & Yamamoto K. (2013). Bifidobacterial alpha-galactosidase with unique carbohydrate-binding module specifically acts on blood group B antigen. *Glycobiology* **23**, 232-240.

Wells, J. M., Wilson, P. W. & Le Page, R. W. (1993). Improved cloning vectors and transformation procedure for *Lactococcus lactis*. *The Journal of Applied Bacteriology* **74**, 629-636.

Wool, D., Hendrix, D. L. & Shukry, O. (2006). Seasonal variation in honeydew sugar content of galling aphids (Aphidoidea: Pemphigidae: Fordinae) feeding on Pistacia: Host ecology and aphid physiology. *Basic and Applied Ecology* **7**, 141-151.

Yazawa, K., Imai, K. & Tamura, Z. (1978). Oligosaccharides and polysaccharides specifically utilisable by bifidobacteria. *Chemical & Pharmaceutical Bulletin* **26**, 3306-3311.

Zhao, H., Lu, L., Xiao, M., Wang, Q., Lu, Y., Liu, C., Wang, P., Kumagai, H. & Yamamoto, K. (2008). Cloning and characterisation of a novel α-galactosidase from *Bifidobacterium breve* 203 capable of synthesising Gal-α-1,4 linkage. *FEMS Microbiology Letters* **285**, 278-283.

Zomer, A., Fernandez, M., Kearney, B., Fitzgerald, G. F., Ventura, M. & van Sinderen, D. (2009). An Interactive Regulatory Network Controls Stress Response in *Bifidobacterium breve* UCC2003. *Journal of Bacteriology* **191**, 7039-7049.

Chapter IV

Transcriptional regulation of the raffinose and melezitose utilisation gene clusters of *Bifidobacterium breve* UCC2003.

4.1 ABSTRACT

Members of the genus *Bifidobacterium* are commonly found in the gastrointestinal tract of mammals, including humans, where their growth is presumed to be dependent on various diet- and/or host-derived carbohydrates. To understand transcriptional control of bifidobacterial carbohydrate metabolism, we investigated two genetic carbohydrate utilisation clusters dedicated to the metabolism of raffinose-type sugars and melezitose. Transcriptomic and gene inactivation approaches revealed that the raffinose utilisation system is positively regulated by an activator protein, designated RafR, while the gene cluster associated with melezitose metabolism is directly negatively regulated by a LacI-type transcriptional regulator, designated MelR1, while there appears to be indirect control by means of a second LacI-type regulator, MelR2. *In silico* analysis, DNA/protein interaction and primer extension studies revealed the MelR1 and MelR2 operator sequences, each of which is positioned just upstream of, or overlapping, the correspondingly regulated promoter sequences. Similar analyses identified the RafR binding operator sequence located upstream of the *raf*B promoter. Interestingly, we show that melezitose is preferentially utilised over raffinose, indicating a carbohydrate utilisation hierarchy in bifidobacteria with regards to particular oligosaccharides.

4.2 INTRODUCTION

Bifidobacteria are saccharolytic, non-motile, non-sporulating, Gram-positive, anaerobic bacteria, which possess a high GC genome content, and belong to the phylum *Actinobacteria.* Bifidobacteria typically exhibit a forked (bifid) morphology and are commonly found as anaerobic commensals of the gastrointestinal tract (GIT) (Miyake *et al.*, 1998; Ventura *et al.*, 2007). Bifidobacteria were first isolated by Tissier (Tissier, 1900) as natural inhabitants of the gastrointestinal tract of humans and mammals, and he was also the first to recognise their probiotic or health-promoting potential (H.Tissier, 1906). Consistent with Tissiers observations it is known that members of certain bacterial genera, such as bifidobacteria and lactobacilli, provide (some degree of) protection against certain types of gastrointestinal infections (Erdo *et al.*, 2012; Parker *et al.*, 2013).

It is also becoming increasingly apparent that a 'normal' (or healthy) colonic microbiota may reduce the risk and/or pathology of gastrointestinal diseases and disorders, such as ulcerative colitis, bowel cancer and pseudomembranous colitis (Gough *et al.*, 2011; Rohlke & Stollman, 2012; Zhu *et al.*, 2013). Several recent studies have focused on the use of prebiotics, which are food ingredients that beneficially affect the host by selectively stimulating growth and/or activity of one or a limited number of beneficial bacteria in the colon (Gibson & Roberfroid, 1995), thereby improving host health. Several carbohydrates have been shown to exert prebiotic effects and include whole grain wheat, fructo-oligosaccharides, galacto-oligosaccharides and lactulose (Bouhnik *et al.*, 1999; Bouhnik *et al.*, 2004; Costabile *et al.*, 2008; Macfarlane *et al.*, 2008).

Metagenomic analyses of the human gut microbiota have generated knowledge that may allow for the rational selection of novel prebiotics that maintain and/or enhance a healthy gut microbiota (Cecchini *et al.*, 2013). Integral to the selection of carbohydrates with prebiotic potential is an in-depth understanding of carbohydrate metabolism by members of the gut microbiota. In this respect bifidobacteria possess a unique hexose metabolism, the so-called bifid-shunt, a metabolic pathway that employs the signature enzyme fructose-6-Phosphate phosphoketolase (F6PPK), converting carbohydrates to mainly acetic and lactic acid (Scardovi & Trovatelli, 1965).

Our recent work has established that *Bifidobacterium breve* UCC2003, an isolate from nursling stool, can metabolise a remarkable range of mono-, di-, oligo- and polysaccharides (Maze *et al.*, 2007; O'Connell *et al.*, 2013; O'Connell Motherway *et al.*, 2008; O'Connell Motherway *et al.*, 2011a; Pokusaeva *et al.*, 2010; Pokusaeva *et al.*, 2011; Ryan *et al.*, 2005). Transcription of bifidobacterial gene clusters associated with carbohydrate metabolism appears to be tightly regulated and several observations have been made suggesting that such clusters are subject to carbon catabolite control (Parche *et al.*, 2006).

Direct transcriptional control of various gene clusters involved in bifidobacterial carbohydrate metabolism is achieved by LacI-type repressors, representing LacI/GalR family proteins that typically consist of an N-terminal helix–turn–helix (HTH) DNA-binding motif and a C-terminal domain for oligomerisation and effector binding (Fukami-Kobayashi *et al.*, 2003; Weickert & Adhya, 1992). In contrast to other bacteria, the number of (predicted) phosphoenolpyruvate-phosphotransferase systems (PEP-PTSs) encoded by bifidobacteria is relatively low (Maze *et al.*, 2007; Parche *et al.*, 2006). Specific elements of the (glucose-specific) PEP-PTS system are

in many bacteria responsible for carbon catabolite control (Deutscher, 2008), however, this does not seem to be case in bifidobacteria as the required regulatory proteins appear to be absent. Therefore, the actual manner by which catabolite repression, if operational, is achieved by bifidobacteria is currently unknown.

Recently, we have identified a *raf* gene cluster dedicated to the metabolism of the plant-derived melibiose/raffinose-containing carbohydrates in *B. breve* UCC2003 (O'Connell *et al.*, 2013). The *raf* locus in the *B. breve* UCC2003 genome is located adjacent to the *mel* gene cluster, which allows this strain to metabolise the trisaccharide melezitose [α-D-Glc*p*-(1→3)-β-D-Fru*f*-(2→1)-α-D-Glc*p*], found in honeydew and manna (Bacon & Dickinson, 1957; O'Connell *et al.*, 2013). In the current work the aim was to investigate the transcriptional regulation of the gene clusters associated with the metabolism of the raffinose family sugars and melezitose.

4.3 MATERIALS AND METHODS

Bacterial strains, plasmids and culture conditions. Bacterial strains and plasmids used in this study are detailed in Table 4.1. Bifidobacteria were routinely cultured in either de Mann Rogosa and Sharpe medium (MRS; Difco™, BD, Le Pont de Claix, France), supplemented with 0.05 % cysteine-HCl, or reinforced clostridial medium (RCM; Oxoid Ltd.). Carbohydrate utilisation by bifidobacterial strains was examined in modified de Mann Rogosa and Sharpe Medium (mMRS) prepared from first principles (De Man *et al.*, 1960), and to which, prior to inoculation, cysteine-HCl (0.05 % w/v) and a particular carbohydrate source (1 % w/v) had been added. The carbohydrates used were raffinose, stachyose, melibiose, melezitose and lactose (all purchased from Sigma-Aldrich, Steinheim, Germany). Bifidobacterial cultures were incubated at 37^{o}C under anaerobic conditions which were maintained using an Anaerocult oxygen depleting system (Merck, Darmstadt, Germany) in an anaerobic chamber. *Lactococcus lactis* strains were cultivated in M17 broth containing 0.5 % glucose (Terzaghi & Sandine, 1975) at 30^{o}C. *Escherichia coli* strains were cultured in Luria-Bertani broth (LB) (Sambrook *et al.*, 1989) at 37^{o}C with agitation. Where appropriate, growth media contained chloramphenicol (Cm; 5 µg ml^{-1} for *L. lactis*, 10 µg ml^{-1} for *E. coli*, and 2.5 µg ml^{-1} for *B. breve*), erythromycin (Em; 100 µg ml^{-1} for *E. coli*), tetracycline (Tet; 10 µg ml^{-1} for *E. coli* or *B. breve)*, kanamycin (Km; 50 µg ml^{-1} for *E. coli*) or Ampicillin (Amp; 100 µg ml^{-1} for *E. coli*).

In order to determine bacterial growth profiles and final optical densities, five millilitres of freshly prepared mMRS medium including a particular carbohydrate (see above) was inoculated with 50 µl (1 %) of a stationary phase culture of a particular *B. breve* strain. Un-inoculated mMRS was used as a negative control. Cultures were incubated anaerobically at 37^{o}C for 16 hr, and the optical density at

600 nm (OD_{600nm}) was determined at 30 min intervals using a Powerwave™ microplate spectrophotometer (BioTek Instruments, Inc. USA) in conjunction with Gen5™ Microplate Software for Windows.

Nucleotide sequence analysis. Sequence data was obtained from the Artemis-mediated (Rutherford *et al.*, 2000) genome annotations of *B. breve* UCC2003 (O'Connell Motherway *et al.*, 2011b). Database searches were performed using non-redundant sequences accessible at the National Centre for Biotechnology Information (http://www.ncbi.nlm.nih.gov) using Blast (Altschul *et al.*, 1990; Altschul *et al.*, 1997). Sequence verification and analysis were performed using the Seqman and Seqbuilder programs of the DNASTAR software package (DNASTAR, Madison, WI, USA v10.1.2).

DNA manipulations. Chromosomal DNA was isolated from bifidobacteria as previously described (O'Riordan & Fitzgerald, 1999). Minipreparation of plasmid DNA from *E. coli* or *L. lactis* was achieved using the Qiaprep spin plasmid miniprep kit (Qiagen GmBH, Hilden, Germany). For *L. lactis* an initial lysis step was incorporated into the plasmid isolation procedure by resuspending cells in lysis buffer supplemented with lysozyme (30 mg ml^{-1}) followed by incubation at 37^{o}C for 30 min (Sambrook *et al.*, 1989). Procedures for DNA manipulations were performed essentially as described previously. Restriction enzymes and T4 DNA ligase were used according to the supplier's instructions (Roche Diagnostics, Bell Lane, East Sussex, UK). Synthetic single stranded oligonucleotide primers used in this study, detailed in Table 4.S1, were synthesised by Eurofins (Ebersberg, Germany). Standard PCRs were performed using TaqPCR mastermix (Qiagen), while *B. breve* colony PCRs were performed as described previously (O'Connell Motherway *et al.*, 2009). PCR fragments were purified using the Qiagen PCR purification kit (Qiagen).

Electroporation of plasmid DNA into *E. coli* was performed as previously described (Sambrook *et al.*, 1989). Electrotransformation of *B. breve* UCC2003 (Maze *et al.*, 2007) and *L. lactis* (Wells *et al.*, 1993) was performed according to published protocols. The correct orientation and integrity of all constructs was verified by DNA sequencing, performed by Eurofins (Ebersberg, Germany).

Construction of plasmids pNZ-MelR1-His, pNZ-MelR2-His and pQE30-RafR. DNA fragments containing the complete coding regions of *Bbr_1864* (designated here as *melR1*) and *Bbr_1863* (designated here as *melR2*) were generated by PCR amplification employing chromosomal DNA of *B. breve* UCC2003 as a template, employing PFU DNA polymerase (Agilent technologies) and primer combinations melR1EcorVF and melR1Xba1R, and melR2Nco1F and melR2Xba1R, respectively (Table 4.S1). NcoI or EcoRV, and XbaI restriction sites were incorporated at the 5' ends of each forward and reverse primer combination, respectively, for the *melR1* and *melR2*-encompassing primers (Table 4.S1). In addition, an in frame His_{10}-encoding sequence was incorporated into each of the forward primers to facilitate downstream protein purification using the Ni-NTA affinity system (Qiagen). The *melR1* and *melR2*-encompassing amplicons were digested with NcoI/EcoRV and XbaI, and ligated into the NcoI/ScaI and XbaI-digested nisin-inducible translational fusion plasmids pNZ8048 or pNZ8150, respectively (Mierau *et al.*, 2005). Ligation mixtures were introduced into *L. lactis* NZ9000 (Table 4.1) by electrotransformation, and transformants were then selected based on chloramphenicol resistance. The plasmid content of a number of Cm^r transformants was screened by restriction analysis and the integrity of positively identified recombinant *melR1* or *melR2*-containing plasmids was verified by sequencing, resulting in constructs pNZ-MelR1-His or pNZ-MelR2-His, respectively.

The coding region of the predicted ROK-type regulator-encoding gene *Bbr_1868* (designated here as *rafR*) was PCR-amplified employing Taq DNA polymerase and chromosomal DNA of *B. breve* UCC2003 as a template, and primer combinations ROKBglIIF and ROKPst1R, which had BglII and Pst1 restriction sites incorporated at the 5' ends of the forward and reverse primers, respectively (Table 4.S1). The generated *rafR*-encompassing fragment was digested with BglII and Pst1, and ligated to similarly digested pQE30 (Qiagen), an IPTG-inducible protein expression plasmid, which allows incorporation of a His-tag into the N-terminus of the expressed protein target. The ligation mixture was introduced into *E. coli* XL1 blue (Table 4.1) by electrotransformation, and transformants were selected based on ampicillin and tetracycline resistance. The plasmid content of a number of Amp^r and Tet^r transformants was screened by restriction analysis and the integrity of a positively identified recombinant plasmid was verified by sequencing and designated pQE30-RafR.

Protein (over)production and purification. For the (over)production of MelR1 and MelR2 25 ml of M17 broth supplemented with 0.5 % glucose was inoculated with a 2 % inoculum of *L. lactis* NZ9000 harbouring either pNZ-MelR1-His, pNZ-MelR2-His or the empty vector pNZ8048 (used as a negative control), followed by incubation at 30^{o}C until an Optical Density (OD_{600nm}) of 0.5 was reached, at which point protein expression was induced by the addition of filter sterilised nisin-containing, cell free supernatant of *L. lactis* NZ9700 followed by continued incubation at 30^{o}C for 90 minutes. Cells were harvested, resuspended in 10 mM Tris-HCl buffer (pH 7.0), and disrupted with glass beads in a mini-bead beater (BioSpec Products, Bartlesville, OK). Cellular debris was removed by centrifugation at 3,000g to produce a crude cell extract. Although protein purification of MelR1-His

and MelR2-His was achieved using His-tag affinity chromatography, the purification procedure appeared to render the proteins inactive in subsequent EMSA experiments (results not shown). For this reason crude cell extracts, prepared in a 10 mM Tris-HCl lysis buffer (pH 7.0), were adopted for the EMSA experiments (see below).

In the case of RafR, a 300 ml volume of LB supplemented with tetracycline and ampicillin was inoculated with 6 ml of an overnight culture of *E. coli* XL1-Blue cells harbouring pQE30-RafR, and then incubated at 37°C. At an OD_{600} of 0.5, expression of RafR was induced by the addition of 1 mM IPTG (Roche Diagnostics Ltd., West Sussex, United Kingdom). Following 2 h of incubation in the presence of IPTG, cells were harvested by centrifugation and crude cell extracts for protein purification were prepared as described above. Subsequent protein purification was performed using a PrepEase® kit for His-tagged protein purification (USB, Germany). Elution fractions were analysed by SDS polyacrylamide gel electrophoresis, as described previously (Laemmli, 1970) on a 12.5 % polyacrylamide gel. Following electrophoresis the gels were fixed and stained with Coomassie Brilliant blue to identify fractions containing the purified protein. Prestained Rainbow low molecular weight protein marker (New England Biolabs, Herdfordshire, UK) was used to estimate the molecular weight of the expressed and/or purified proteins.

Construction of B. breve insertion mutant strains. Internal fragments of *rafR* (528 bp; representing codons 88 through to 264 out of the 403 codons of this gene), *melR1* (380 bp; representing codons 107 through to 234 out of the 340 codons of this gene), *melR2* (402 bp; representing codons 93 through to 227 out of the 343 codons of this gene), were amplified by PCR using *B. breve* UCC2003 chromosomal DNA as template and the oligonucleotide primer combinations rafRfHd3in and rafRrxbaIin, melR1fHd3in and melR1rxba1in, and melR2fHd3in and melR2rxba1in,

respectively (Table S4.1). The generated PCR products were ligated to pORI19, an Ori+ RepA- integration plasmid (Law *et al.*, 1995), using HindIII and XbaI restriction sites that were incorporated into the primers for the partial *rafR, melR1* and *melR2*-encompassing amplicons, and introduced into *E. coli* EC101 by electroporation. Recombinant *E. coli* EC101 derivatives containing pORI19 constructs were selected on LB agar containing Em, and supplemented with X-gal (5-bromo-4-chloro-3-indolyl-D-galactopyranoside) (40 μg ml^{-1}) and 1 mM IPTG.

The expected genetic structure of the recombinant plasmids, pORI19-rafR (pORI19 containing an internal 528 bp fragment of the *rafR* gene), pORI19-melR1 (pORI19 containing an internal 402 bp fragment of the *melR1* gene) and pORI19-melR2 (pORI19 containing an internal 380 bp fragment of the *melR2* gene), was confirmed by restriction mapping prior to subcloning of the Tet resistance antibiotic cassette, *tetW*, from pAM5 (Alvarez-Martin *et al.*, 2007) as a SacI fragment into the unique SacI site present on each of the pORI19 derivatives. The orientation of the tetracycline resistance gene in each of the resulting plasmids pORI19-tet-rafR, pORI19-tet-melR1 and pORI19-tet-melR2, was determined by restriction analysis. The plasmids were subsequently introduced into *E. coli* EC101 pNZ-MBbrI-MBbrII (O'Connell Motherway *et al.*, 2009) (transformants were selected based on Cm and Tet resistance) in order to methylate the plasmid constructs before introduction into *B. breve* UCC2003.

Methylation of the plasmid complement of the obtained transformants in EC101 pNZ-MBbrI-MBbrII was confirmed by their observed resistance to PstI restriction (O'Connell Motherway *et al.*, 2009). Plasmid preparations of methylated pORI19-tet-rafR, pORI19-tet-melR1 and pORI19-tet-melR2 were introduced by electroporation into *B. breve* UCC2003 with subsequent selection on RCA plates

supplemented with Tet. Insertion mutants resulting from site-specific homologous recombination were initially confirmed by colony PCR targeting the tetracycline resistance gene *tetW,* followed by a second confirmatory PCR adopting a *tetW*-based primer, either forward or reverse depending on the orientation of *tetW,* in combination with a primer specific for each targeted gene to confirm integration at the correct chromosomal position (Table S4.1). In this case a product of a particular size would only be obtained if the correct gene disruption had been achieved. In this manner mutants were obtained carrying chromosomal insertions in either the *rafR, melR1* or *melR2* gene, resulting in strains *B. breve* UCC2003-*rafR*, *B. breve* UCC2003-*melR1* and *B. breve* UCC2003-*melR2*, respectively.

Analysis of preferential carbohydrate (raffinose ot melezitose) consumption by B. breve UCC2003. Filter-sterilised raffinose and melezitose solutions were added to mMRS medium (to obtain a final sugar concentration of 0.5 % and/or 0.1 % respectively). *B. breve* UCC2003 growth in mMRS supplemented with raffinose/melezitose and 0.05 % (wt/vol) cysteine-HCl was monitored every two hours for 13 hours followed by 2 time points at 24 and 48 hours by measuring the optical density at 600 nm (OD_{600}). In addition, HPAEC-PAD analysis of heat treated (80^{o}C for 2 mins) filtered, cell-free growth medium samples taken at the same time points was performed to analyse raffinose and/or melezitose consumption.

Analysis of global gene expression using B. breve DNA microarrays. Global gene transcription was determined by microarray analysis during growth of *B. breve* UCC2003-*rafR, B. breve* UCC2003-*melR1* and *B. breve* UCC2003-*melR2*, in mMRS supplemented with ribose and these transcriptomes were compared to that obtained from *B. breve* UCC2003 wild type cells that also had been grown on ribose as the

sole carbohydrate source. DNA-microarrays containing oligonucleotide primers representing each of the 1864 annotated genes on the genome of *B. breve* UCC2003 were designed by and obtained from Agilent Technologies (Palo Alto, Ca., USA). Methods for cell disruption, RNA isolation, RNA quality control, complementary DNA synthesis and labelling were performed as described previously (Zomer *et al.*, 2009). Labelled cDNA was hybridised using the Agilent Gene Expression hybridisation kit (part number 5188-5242) as described in the Agilent Two-Color Microarray-Based Gene Expression Analysis v4.0 manual (G4140-90050). Following hybridisation, microarrays were washed in accordance with Agilent's standard procedures and scanned using an Agilent DNA microarray scanner (model G2565A). Generated scans were converted to data files with Agilent's Feature Extraction software (Version 9.5). DNA-microarray data were processed as previously described (Garcia De La Nava *et al.*, 2003; van Hijum *et al.*, 2003; van Hijum *et al.*, 2005). Differential expression tests were performed with the Cyber-T implementation of a variant of the *t*-test (Long *et al.*, 2001). A gene was considered to exhibit significantly differential expression when $p < 0.001$, and an expression ratio of >3 or <0.33 relative to the control. Obtained microarray data were deposited in NCBI's Gene Expression Omnibus and are accessible through GEO Series accession number GSE50211.

Electrophoretic mobility shift assay (EMSA). DNA fragments representing different portions of the promoter regions upstream the *rafA*, *rafB*, *Bbr_1862* and *melA* genes were prepared by PCR using IRD800-labelled primers pairs (MWG Biotech; supplemental Table 4.S2). EMSAs were performed essentially as described previously (Hellman & Fried, 2007**).** In all cases, the binding reactions were performed in a final volume of 20 µl in the presence of poly[d(I-C)] in binding

buffer (20 mM Tris-HCl, 5 mM $MgCl_2$, 0.5 mM DTT, 1 mM EDTA, 200 mM KCl, 10 % glycerol at pH 7.0). Varying amounts of a crude cell extract containing either MelR1 or MelR2, or purified RafR, were mixed on ice with a fixed amount of DNA probe (0.1 pmol) and subsequently incubated for 30 min at 37°C. Samples were loaded onto a 6 % non-denaturing PAA gel prepared in TAE buffer (40 mM Tris acetate [pH 8.0], 2 mM EDTA) and run in a 0.5-to-2.0 x gradient of TAE at 100 V for 90 minutes in an Atto Mini PAGE system (Atto Bioscience and Biotechnology, Tokyo, Japan). Signals were detected using an Odyssey Infrared Imaging System (Li-Cor Biosciences UK Ltd, Cambridge, UK) and images were captured using the supplied Odyssey software V3.0.

Primer extension analysis. Total RNA was isolated from exponentially growing cells of *B. breve* UCC2003 or *B. breve* UCC2003-*melR2*, cultured in mMRS supplemented with 1 % raffinose or melezitose for *B. breve* UCC2003, or in mMRS supplemented with 1 % ribose for *B. breve* UCC2003-*melR2* (Kuipers *et al.*, 1993). Primer extension was performed by annealing 1 pmol of an IRD800-labeled synthetic oligonucleotide to 20 µg of RNA as previously described (Ventura *et al.*, 2005) using primers rafAPERP1, rafAPERP2, rafBPERP1, rafBPERP2, melAPERP1, melAPERP1, 1862PERP1, 1862PERP1(Table 4.S2). Sequence ladders of the presumed *rafA, rafB, melA* and *Bbr_1862* promoter regions were produced using the same primer as the primer extension reaction and employing the DNA cycle Sequencing Kit (Jena Bioscience, Loebstedter Strasse 80, Germany) and were run alongside the primer extension products to allow precise positioning of the transcriptional start site. Separation was achieved on a 6.5 % Li-Cor Matrix KB Plus acrylamide gel. Signal detection and image capture was performed by means of a Li-cor sequencing instrument (Li-Cor Biosciences).

4.4 RESULTS AND DISCUSSION

Metabolism of melezitose in B. breve UCC2003 is preferred over that of raffinose-like carbohydrates

In a previous study (O'Connell *et al.*, 2013) we investigated the metabolism of raffinose-related carbohydrates, raffinose, stachyose and melibiose by *B. breve* UCC2003. This particular study allowed the identification of a gene cluster, comprising *rafA, rafR, rafB, rafC* and *rafD*, predicted to be involved in the utilisation of raffinose-related sugars, and where *rafA* was shown to encode an α-galactosidase, *rafR* a putative ROK-type transcriptional regulator, while the adjacent *rafBCD* genes specified a putative ABC-type sugar uptake system (Fig. 4.1). This same study also identified a gene cluster in the *B. breve* UCC2003 genome, located in the immediate vicinity of the *raf* gene cluster, dedicated to melezitose metabolism (O'Connell *et al.*, 2013), and comprised of the *melA, melB, melC, melD* and *melE*, genes, with *melABC* specifying a presumed ABC-type sugar uptake system, while *melD* and *melE* encoding an α-glucosidase and an α-galactosidase, respectively (Fig. 4.1). Located between these afore mentioned *raf* and *mel* genes are four genes, two of which, designated *melR1* and *melR2,* encode putative LacI-type transcriptional regulators (Fig. 4.1).

We previously (O'Connell *et al.*, 2013) demonstrated that transcription of the *melABCDE* and *rafABCD* genes is induced by the presence of melezitose and raffinose (or raffinose-like sugars), indicative of a substrate-specific transcriptional regulation. In order to investigate if *B. breve* UCC2003 exhibits a preference for either melezitose or raffinose, we determined growth profiles of this strain in the presence of various amounts of melezitose, raffinose or a combination of these two sugars, while we also determined the carbohydrate utilisation from the growth medium by HPAEC-PAD analysis (see Materials and Methods). Following initial

experiments required to determine the precise amounts of carbohydrate needed to clearly visualise preferential sugar utilisation (data not shown), it was found that when *B. breve* UCC2003 was grown in a combination of 0.1 % melezitose and 0.5 % raffinose, the former carbohydrate was utilised first, and that raffinose was being used only when melezitose had been exhausted from the medium (Fig. 4.2A). A diauxic growth curve could be observed when UCC2003 was grown on this melezitose/raffinose-containing medium, although the growth rate differences between melezitose and raffinose-supported growth were negligible under the growth conditions used (Fig.4.2B). These findings indicate that the genes involved in raffinose metabolism are not transcribed in the presence of melezitose, reminiscent of carbon catabolite regulatory control.

Transcriptomes of B. breve UCC2003 mutants carrying disruptions in melR1, melR2 or rafR.

In order to identify and investigate the regulatory players that control the raffinose and melezitose-induced genes, insertion mutants were made in the *melR1*, *melR2* and *rafR* genes that were located within the *mel-raf* region and predicted to encode regulatory proteins (see above and Fig. 4.1), resulting in strains *B. breve* UCC2003-*melR1*, *B. breve* UCC2003-*melR2* and UCC2003-*rafR*, respectively (see Materials and Methods). It was hoped that if any of the putative regulatory genes encoded a repressor, mutation of the gene would lead to increased transcription even in the absence of the inducing carbohydrate.

Altered gene expression patterns were determined by DNA microarray analysis from two independent biological replicates and employing *B. breve* UCC2003, *B. breve* UCC2003-*melR1*, *B. breve* UCC2003-*melR2* or UCC2003-*rafR* cultures grown to the exponential growth phase in mMRS supplemented with ribose. The obtained

transcriptomes revealed that, compared to *B. breve* UCC2003, the *melABCDE* genes were significantly upregulated (fold change > 5.0, P<0.001) in *B. breve* UCC2003-*melR1*, while *melABCDE* as well as its adjacent genes *Bbr_1861* and *Bbr_1862* were upregulated (fold change > 5.0, P<0.001) in *B. breve* UCC2003-*melR2* (Table 4.2). However, the insertion mutation in *melR2* appears to have a polar effect on the transcription of the downstream located gene, *mel*R1, as this gene exhibited a 1.7-fold lower level of transcription in the *B. breve* UCC2003-*melR2* mutant (compared to UCC2003; P<0.001), thereby indicating that *mel*R2 and *mel*R1 are co-transcribed. The obtained data is consistent with our prediction that both LacI-type regulators act as typical transcriptional repressors since the mutations in *B. breve* UCC2003-*melR1* and UCC2003-*melR2* cause transcriptional upregulation of adjacent genes, several of which have been associated with melezitose metabolism. However, the observed transcriptional increases observed for UCC2003-*melR2* may be partly due to a polar effect of the *melR2* mutation on *melR1* transcription.

The transcriptome of *B. breve* UCC2003-*rafR* did not reveal any statistically significant differences with that of *B. breve* UCC2003 that could be linked to the metabolism of raffinose-like sugars (results not shown), which is consistent with the notion that RafR is a ROK-type transcriptional activator. The three mutants were also investigated for their ability to utilise melezitose or raffinose-like sugars. The *B. breve* UCC2003-*melR1* and UCC2003-*melR2* mutants both retained the ability to utilise melezitose as sole carbohydrate source (Fig. 4.3A). Interestingly, the *B. breve* UCC2003-*rafR* was incapable of growing to a high optical density in mMRS supplemented with any of the raffinose-like sugars as the sole carbon source (Fig. 4.3B), while growth of *B. breve* UCC2003-*rafR* in mMRS supplemented with 1 % lactose was comparable to that observed for the parent strain *B. breve* UCC2003.

These results demonstrate that *rafR* is required for growth on raffinose-related carbohydrates and suggest that the ROK-type transcriptional regulator RafA is required for the activation of the raffinose gene cluster.

MelR1 and MelR2 bind to the melA and Bbr1862 promoter regions, respectively

The transcriptome data obtained for the *B. breve* UCC2003-*melR1* and UCC2003-*melR2* mutants (see above) suggests that *melR1* and *melR2* encode LacI-type regulators that control various genes of the *mel* gene cluster. In order to establish if the MelR1 or MelR2 proteins directly and specifically interact with presumed operator sequences within the promoter region(s) of the *mel* gene cluster, i.e. the *Bbr_1862* and/or *melA* promoter regions (see below), electrophoretic mobility shift assays (EMSAs) were performed. For this purpose, we first cloned *melR1* and *melR2* in the nisin-inducible vectors pNZ8048 and pNZ8150, respectively, with the introduction of a N-terminal His-tag-encoding sequence to facilitate protein expression and purification in *L. lactis* NZ9000 (and generating plasmids pNZ-MelR1-His and pNZ-MelR2-His, respectively).

However, although MelR1and MelR2 could be obtained as purified proteins, they proved unable to form DNA-protein complexes under the conditions used, which had been noted previously for LacI-type regulators in bifidobacteria (Pokusaeva *et al.*, 2011). For this reason, instead of the purified proteins, crude cell extracts of (nisin-induced) *L. lactis* NZ9000 pNZ-MelR1-His or *L. lactis* NZ9000 pNZ-MelR2-His were employed to perform EMSAs. As a negative control for EMSAs we used crude cell extract of *L. lactis* NZ9000 (pNZ8048) (empty vector) incubated with various DNA fragments as mentioned below, which, as expected, all failed to alter electrophoretic behaviour of these DNA fragments. The obtained results with a crude cell extract containing MelR1 clearly demonstrate that this protein binds to the

IRD800-labelled DNA fragment M1, which encompasses the 444 bp promoter-containing (see below) region preceding *melA*, while MelR1 failed to bind to the IRD800-labelled DNA fragment K1, encompassing the *Bbr_1862* promoter region, or the *rafA* or *rafB* promoter regions (Fig. 4.4A and 4.4 B (i), and results not shown).

Further dissection of the *melA* promoter region showed that MelR1 binding required an 84 bp DNA segment which is present within fragments M1- M4 [Fig. 4.4B (i) and Supplementary figure S4.1 (i-iii)], and which contains a 24 bp imperfect inverted repeat, 5'-TCATGCATAAGC><GCTTAGCAAATA-3'representing the putative MelR1-specific operator sequence. This was further validated by EMSAs using a 46 bp synthetic DNA fragment that contained this MelR1 binding (motifmelAIR) (Fig. 4.4C (i)). Furthermore, introduction of two point mutations in this presumed MelR1 operator sequence (a G to T and a C to A mutation at positions eight and nine of the sequence as indicated above) were shown to prevent binding of MelR1 (mu-melAIR) (Fig. 4.4 C(iii)). Other mutations in bases on either side of this motif did not affect binding of MelR1 (results not shown; for all wild type and mutated primer sequences see Table S4.3). To investigate whether MelR1 interaction with its target DNA sequence is influenced by a carbohydrate effector molecule, as is known for other LacI-type regulators (reviewed by (Swint-Kruse & Matthews, 2009; Wilson *et al.*, 2007)), several carbohydrates were tested for their effects on MelR1–DNA complex formation. The results obtained show that the binding ability of MelR1 to the M1 fragment of the *melA* promoter region is lost in the presence of melezitose at concentrations > 2.5 mM (Fig. 4.4D), whereas under the same experimental conditions galactose or lactose did not affect MelR1 binding to its target (Supplementary figure S4.2 (i) and (ii)).

When EMSAs were performed using a nisin-induced *L. lactis* NZ9000 pNZ-MelR2-His cell extract, MelR2 was shown to bind to IRD800-labelled DNA fragment K1, which encompasses the 310 bp promoter-containing (see below) region preceding *Bbr_1862*, but not with similarly labelled DNA fragments containing the promoter regions upstream of *melA* (fragment M1), *rafA* or *rafB* (Fig. 4.4A and 4.4 B (ii), and results not shown). Further EMSAs showed that MelR2 binding required a 85 bp DNA segment present within fragments K1 and K2, and absent in fragments K3 and K4 of the *Bbr_1862* promoter region (Fig. 4.4B (ii); Supplementary figure S4.3 (i-iii). Inspection of this 85 bp fragment revealed the presence of a 24 bp imperfect inverted repeat, 5'-ATGTGCGTAATC><GATATCGCAAAT-3', representing a putative MelR2 operator sequence. This was validated by EMSAs using a synthetic 50 bp DNA fragment that contained this predicted operator sequence (Bbr1862IR) (Fig. 4.4C (ii)). Introduction of two point mutations in the putative 24 bp MelR2 binding motif (an A to C and a T to G mutation at positions ten and eleven of this motif) were shown to prevent MelR2 binding (See Supplementary table S4.4) for IRD800 primer sequences (mu-Bbr1862IR) (Fig. 4.4 C (iv)). Other mutations were also made at positions one and two and also in bases surrounding this motif but none of these mutations was shown to affect binding under the conditions tested (results not shown; see Supplementary table S4.4 for relevant primer sequences used). To investigate if MelR2 interaction with its target DNA sequence is influenced by a carbohydrate effector molecule, as shown for MelR1, several carbohydrates were tested for their effects on MelR2–DNA complex formation. However, none of the carbohydrates tested (melezitose, turanose, sucrose, glucose, fructose, α-(1→3) galactobiose or α-(1→4) galactobiose) was shown to prevent MelR2 binding to its

DNA target (Supplementary figure S4.4 (i-vii)), suggesting that MelR2 binding is abrogated by an as yet unknown carbohydrate (O'Connell *et al.*, 2013).

The above binding studies, combined with the transcriptome analyses indicate that MelR1 is the melezitose-responsive repressor of the *melABCDE* gene cluster, while MelR2 is a LacI-type repressor controlling transcription of *Bbr_1861* and *Bbr_1862*. The observation that the *melR2* mutation leads to an increase in *melABCDE* transcription is most likely an indirect effect due to its negative polar effect on *melR1* transcription, leading to reduced MelR1 expression and consequently derepression of the *melABCDE* gene cluster.

RafR binds to the rafB promoter region The presence of *rafR*, encoding a putative ROK-type regulator within the raffinose gene cluster suggests that this gene plays a role in the transcriptional regulation of the *raf* gene cluster, a notion that was supported by the inability of a *rafR* mutant to support proper growth on raffinose-type sugars (see above). In order to establish if RafR is capable of interacting with specific sequences within the raffinose-induced (O'Connell *et al.*, 2013) promoter regions upstream of *rafA* and *rafB* (see below), we overexpressed and purified RafR as a His-tagged protein (see Materials and Methods). RafR was then used to perform EMSAs, which clearly demonstrate that it causes a mobility shift of IRD800-labelled DNA fragment R1, which includes the 241 bp promoter region upstream of *rafB* (Fig. 4.5A and B (i)). However, under the conditions used RafR was unable to bind to IRD800-labelled DNA fragment G1 which includes the 223 bp promoter region upstream of *rafA* (results not shown).

Further EMSAs involving RafR and fragments R2 and R3 (Fig. 4.6 (i-ii) of the *rafB* promoter region showed that RafR binding requires an 81 bp DNA segment (Fig.

4.5(i) and (ii)), which contains an imperfect inverted repeat, 5'-TTTATTGCGTT>A<ATCACAAAATA-3', thus representing the putative RafR operator sequence. This was validated by EMSAs using a 38 bp DNA fragment that contained this sequence (rafBIR) (Fig. 4.5B (ii)). Introduction of point mutations in a 35 bp sequence encompassing the putative RafR binding motif revealed various positions within this operator sequence that are crucial for RafR binding (Fig. 4.5C; for primer sequences used see Supplementary table S4.5). This operator binding sequence is very similar to the consensus operator sequence determined for two other members of the ROK family proteins, Mlc and NagC, in *E. coli* (Plumbridge, 2001, Titgemeyer *et al.*, 1994).

RafR interaction with its target DNA sequence was not affected by the presence of raffinose, stachyose, melibose, galactose, glucose, sucrose, α-(1→3)-galactobiose or α-(1→4)-galactobiose at concentrations ranging from 2.5 mM to 20 mM (Supplementary figure S4.5 (i-viii)). Binding of RafR to the *rafA* promoter region was also tested in the presence of these sugars; however, this did not result in binding of RafR to the *rafA* promoter region (results not shown).

Identification of the transcription start sites of rafA, rafB, melA and Bbr_1862

From the genetic organisation (Fig. 4.1) and the expression pattern of the *raf* locus (O'Connell *et al.*, 2013) it was deduced that the *raf* cluster contained two raffinose/stachyose/melibiose-inducible promoters: one in front of the *rafA* gene and one in front of the *rafB* gene. In relation to the *mel* cluster potential promoter regions were identified based on *in silico* analysis of the *mel* locus (Fig. 4.1) coupled with transcriptomic data obtained during growth on melezitose (O'Connell *et al.*, 2013), as well as the transcriptomes of UCC2003-*melR1*and UCC2003-*melR2* (see above). It was reasoned that one melezitose-inducible, MelR1-dependent promoter was

located directly upstream of the *melA* coding sequence and one promoter (transcription of which is subject to MelR2 and an as yet unknown inducer molecule) preceding the *Bbr_1862* coding sequence.

In order to determine the transcription start site of these presumed promoters, primer extension analysis was performed using RNA extracted from *B. breve* UCC2003 grown in mMRS containing either 1 % melezitose or raffinose as appropriate. To determine the transcription start site of *Bbr_1862*, *B. breve* UCC2003-*melR2* (whose transcriptome revealed constitutive transcription of *Bbr_1862*) was grown in mMRS supplemented with ribose as the sole carbohydrate source. Single extension products were identified upstream *mel*A, the *Bbr_1862* gene, *rafB*, and *rafA*, and in all cases potential promoter recognition sequences resembling consensus −10 and −35 hexamers could be identified upstream of these transcription start points (Fig. 4.7 A-D). In the case of *melA* and *Bbr_1862*-associated promoters, the MelR1 and MelR2 binding operator sequences overlap the respective −10 sequence, which would support our findings that MelR1 and MelR2 are LacI-type transcriptional repressors. For the *rafB* promoter region the RafR operator sequence is located upstream of the -35 and -10 promoter recognition sequences, a location that is consistent with the notion that RafR functions as a transcriptional activator.

4.5 CONCLUDING REMARKS

The data assembled in this study provide significant information on *B. breve* UCC2003's ability to regulate two adjacent genetic loci dedicated to the utilisation of raffinose-containing carbohydrates and melezitose (O'Connell *et al.*, 2013). We conclude that MelR1 is a LacI-type regulator that acts as a melezitose-responsive repressor of the *melABCDE* gene cluster, while MelR2 is also a LacI-type regulator that acts as a repressor of the *Bbr_1861* and *Bbr_1862* genes, the latter encoding a predicted carbohydrate binding protein of an ABC-type sugar transport system. Because of its vicinity and the finding that *melE* encodes an α-galactosidase, we speculate that MelR2 responds to an oligosaccharide, containing a melezitose backbone to which one or more α-galactose moieties are attached. This α-galactose-containing oligosaccharide is bound by the product of *Bbr_1862* and internalised by the *melABC* uptake system, after which the α-galactose moiety (or moieties) are removed by the action of MelE followed by further degradation by MelD. Unfortunately, no such substrates are currently commercially available and this hypothesis therefore remains to be verified.

In a classic example by (Irwin & Ptashne, 1987) it was observed that mutating the catabolite activator protein (CAP) resulted in lack of transcriptional activation of the *gal* promoter. Similarly, in *Corynebacterium glutamicum* 13032, a deletion in the gene encoding the ROK-type protein, CysR, rendered this strain incapable of utilising sulphate or sulphonates as the sole source of sulphur leading the authors to speculate that CysR functions as a transcriptional activator (Ruckert *et al.*, 2008). Since an insertion in *rafR* led to the inability of the strain to utilise raffinose-like carbohydrates and combined with our binding studies, where RafR, like many transcriptional activators binds upstream of the promoter region (Collado-Vides *et*

al., 1991; Madan Babu & Teichmann, 2003), we conclude that RafR also acts as a transcriptional activator of the *rafBCD* cluster.

Somewhat to our surprise RafR did not bind to the *rafA* promoter, at least not under the conditions used here, although this promoter is transcriptionally induced when *B. breve* UCC2003 is grown in raffinose-like sugars (O'Connell *et al.*, 2013). It may be that RafR requires an effector for binding to this promoter and future research will investigate the transcriptional regulation of *rafA*.

Finally and interestingly, we found that *B. breve* UCC2003 prefers melezitose metabolism over that of raffinose-like sugars, indicating that this bifidobacterial strain engages in some form of carbon catabolite control. Although regulation of melezitose metabolism is established through the LacI-type repressor MelR1, while raffinose-like sugars are metabolised by means of the ROK-protein RafA, our findings clearly indicate that an additional, but so far elusive, carbohydrate detection and control system must be operating in *B. breve* UCC2003 and perhaps other bifidobacteria. Future research efforts will be aimed at identifying and characterising this regulatory mechanism.

4.6 ACKNOWLEDGEMENTS:

This Chapter has emanated from research conducted with the financial support of Science Foundation Ireland (SFI) under Grant Numbers 07/CE/B1368, 08/SRC/B13404 and SFI/12/RC/2273. KJOC was supported by a postgraduate fellowship funded through the Tomar trust, while MOCM is a recipient of a HRB postdoctoral fellowship (Grant no. PDTM/20011/9). We also sincerely thank Dr Stephen Cunningham, Dr. Marian Keane and Prof. Lokesh Joshi of the Alimentary Glycoscience Research Cluster, NUIG, Ireland, for kindly supplying α-1,4 and α-1,3-galactobiose. I would also like to acknowledge Kieran James for technical assistance.

4.7 TABLES AND FIGURES

Table 4.1 Strains and plasmids used in this study

Strain/ plasmid	Relevant characteristics	Reference or source
E.coli strains		
EC101	Cloning host, repA+ kmr	(Law *et al.*, 1995)
EC101-pNZ-M.Bbrll+Bbr111	EC101 harbouring pNZ8048 derivative containing bbrllM and bbrlllM	
XLI blue	RecA1 endA1 gyrA96 thi-1 hsdR17 supE44 relA1 lac [F′ proAB lacIqZΔM15 Tn10 (Tetr)]	Stratagene
L. lactis strains		
NZ9000	MG1363,nisin-inducible overexpression host; pepN::nisRK	(de Ruyter *et al.*, 1996)
NZ9700	Nisin producing strain	(de Ruyter *et al.*, 1996)
B. breve strains		
UCC2003	Isolate from a nursling stool	(Maze *et al.*, 2007)
UCC2003- *melR1*	pORI19-tet-mel1 insertion mutant of UCC2003	This study
UCC2003- *melR2*	pORI19-tet- melR2 insertion mutant of UCC2003	This study
UCC2003-*rafR*	pORI19-tet-rafR insertion mutant of UCC2003	This study
Plasmids		
pORI19	Emr , repA- , ori+ , cloning vector	(Law *et al.*, 1995)
pORI19-tet- *melR1*	Internal 421bp fragment of melR1 and tetW cloned in pORI19	This study
pORI19-tet- *melR*2	Internal 456bp fragment of melR2 and tetW cloned in pORI19	This study
pORI19-tet-rafR	Internal 331bp fragment of rafR and tetW cloned in pORI19	This study
pAM5	pBC1-puC19-Tcr	(Alvarez-Martin *et al.*, 2007)
pNZ8048	Cmr; nisin-inducible translational fusion vector	(de Ruyter *et al.*, 1996)
pNZ8150	Cmr; nisin-inducible translational fusion vector	(Mierau *et al.*, 2005)
PQE30	Ampr; IPTG-inducible vector	Qiagen
pNZ-MelR1-His	MelR1 with his tag cloned downstream of nisin inducible promoter on pNZ8048	This study
pnz-MelR2-His	melR2 with his tag cloned downstream of nisin inducible promoter on pNZ8150	This study
pQE30-RafR	rafR IPTG-inducible vector PQE30	This study

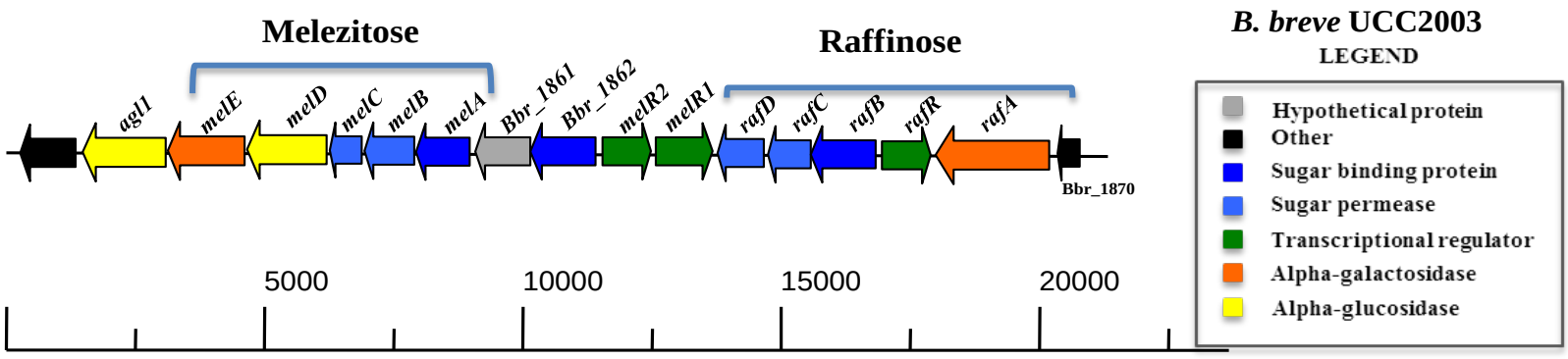

Figure 4.1 Representation of the raffinose and melezitose utilisation operons of *B. breve* UCC2003.

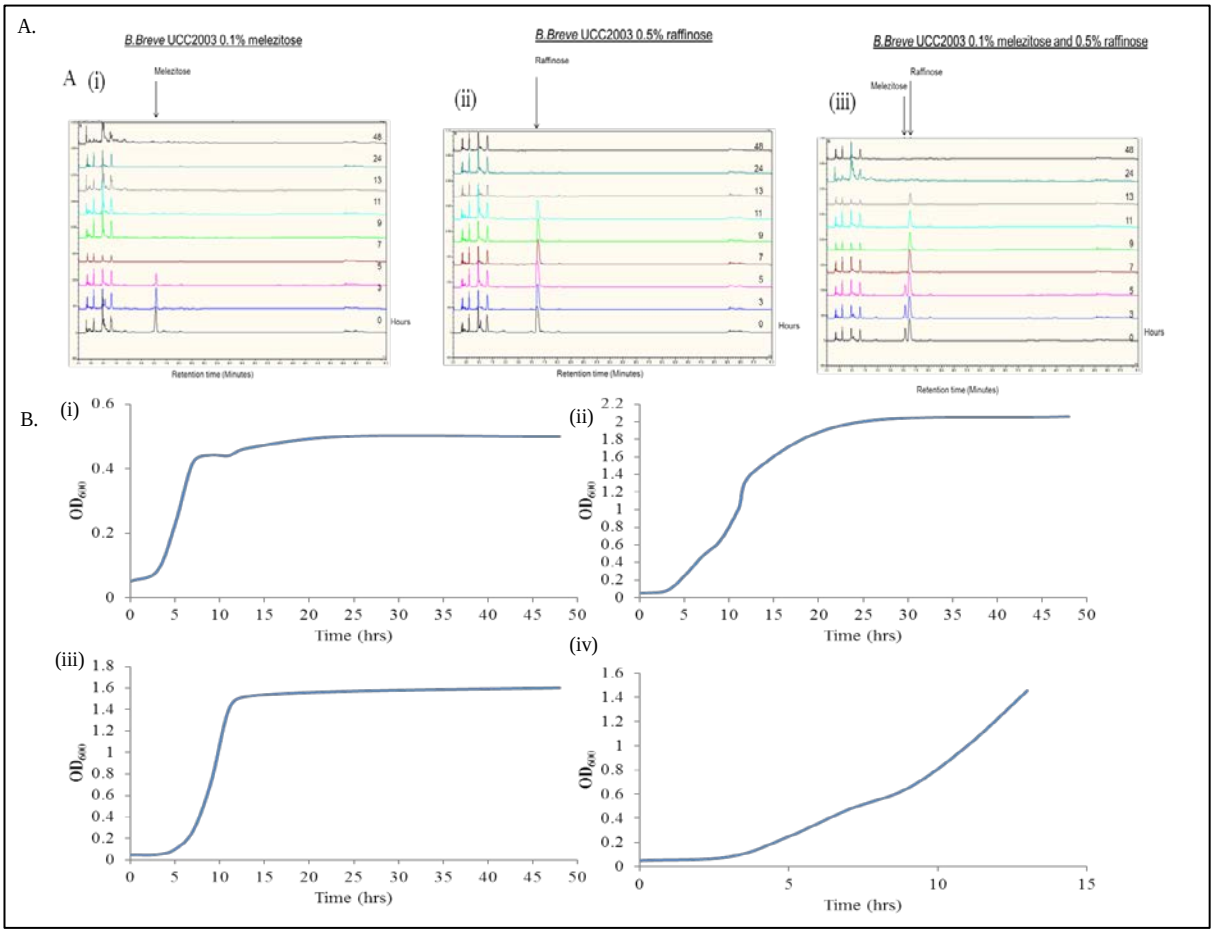

Figure 4.2 A. HPAEC-PAD analysis of *B. breve* UCC2003 consumption of (i) 0.1% melezitose (ii) 0.5% raffinose and (iii) 0.5% raffinose and 0.1% melezitose during 48 hours of growth. B. Consumption of (i) 0.1% melezitose (ii) 0.5% raffinose (iii) mixture 0.5% raffinose and 0.1% melezitose up to 48 hours (iv) mixture 0.5% raffinose and 0.1% melezitose during 15 hours of growth in mMRS medium by *B. breve*UCC2003.

Table 4.2 Transcriptome analysis of *B. breve* UCC2003- *melR1* and *B. breve* UCC2003- *melR2* as compared to *B. breve* UCC2003 grown on 1% ribose.

Gene number	Name	Putative Function	Fold up regulation *melR2*	Fold up regulation *melR1*
Bbr_1856	*melE*	Alpha-galactosidase	6.60	5.51
Bbr_1857	*melD*	Alpha-1,4-glucosidase	10.24	8.59
Bbr_1858	*melC*	Permease protein of ABC transporter system for sugars	86.80	78.10
Bbr_1859	*melB*	Permease protein of ABC transporter system for sugars	88.56	67.90
Bbr_1860	*melA*	Solute binding protein of ABC transporter system for sugars	77.64	80.63
Bbr_1861		Conserved hypothetical protein	59.43	_
Bbr_1862		Solute binding protein of ABC transporter system for sugars	148.62	_

Fold cut off point is 5 fold, p value < 0.001, and values below cut-off are indicated in the Table as a dash.

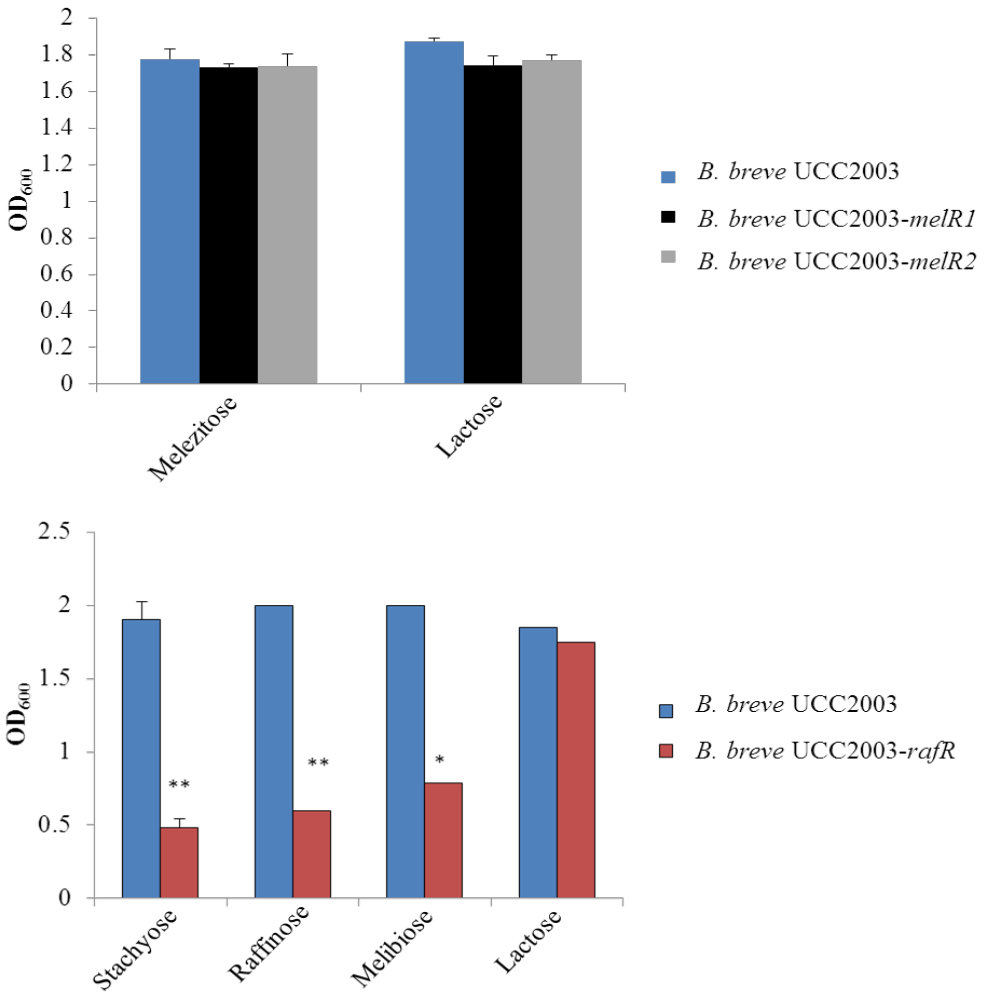

Figure 4.3 A Final OD600 after 16 hours of wild type strain *B. breve* UCC2003 and insertion mutants *B. breve* UCC2003- *melR1* and *B. breve* UCC2003- *melR2* grown on 1% melezitose and 1% lactose. The results are mean values obtained from three separate experiments, $P< 0.1$. B Final OD_{600} after 16 hours of wild type strain *B. breve* UCC2003 and insertion mutant *B. breve* UCC2003-*rafR* grown on 1% stachyose, 1% raffinose, 1% melibiose and 1% lactose (*P<0.01,**P<0.001).

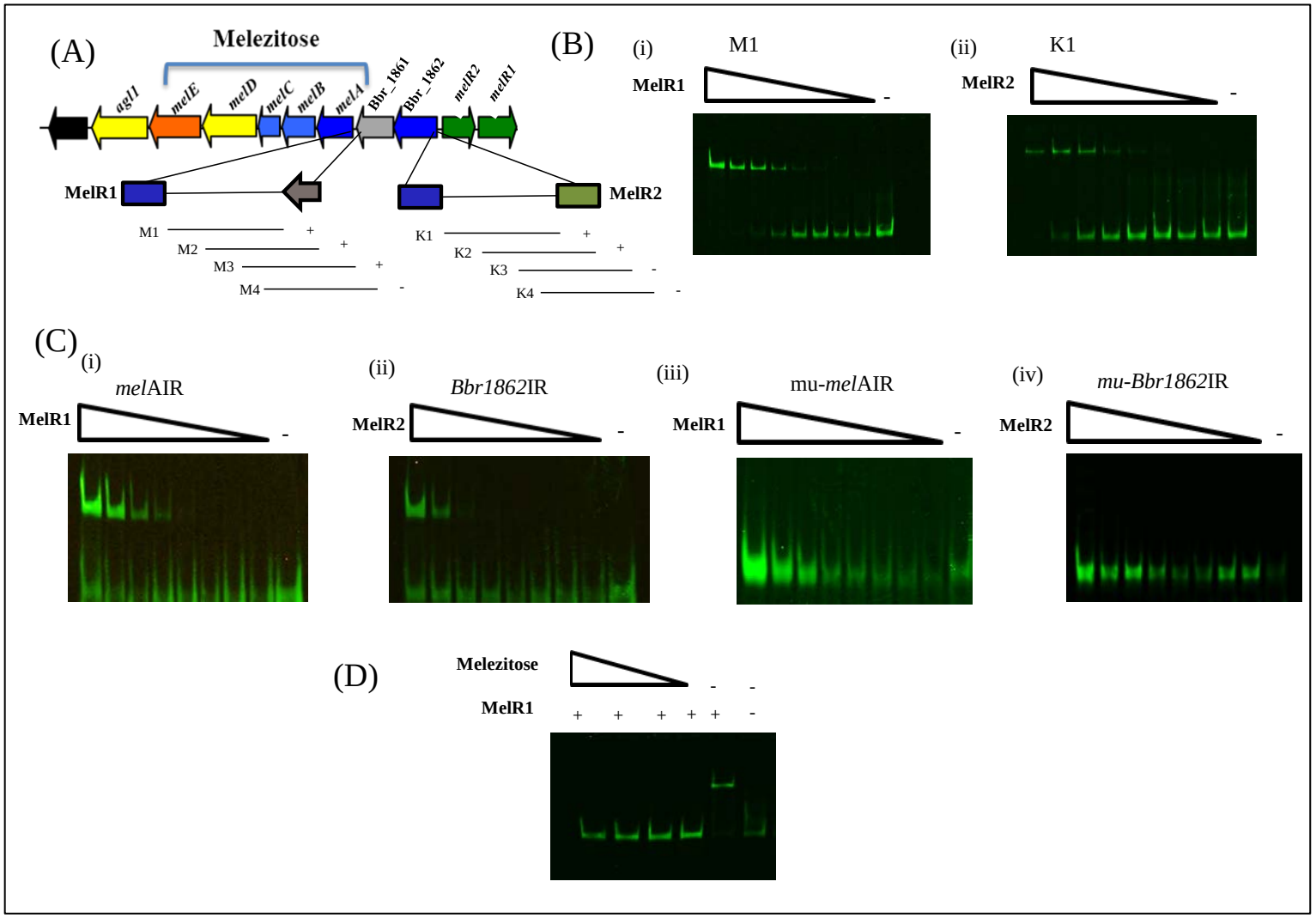
(A)
Melezitose
agl1
melE
melD
melC
melB
melA
Bbr_1861
Bbr_1862
melR2
melR1
MelR1
MelR2
M1
M2
M3
M4
K1
K2
K3
K4
(B)
(i)
M1
MelR1
(ii)
K1
MelR2
(C)
(i)
melAIR
MelR1
(ii)
Bbr1862IR
MelR2
(iii)
mu-melAIR
MelR1
(iv)
mu-Bbr1862IR
MelR2
(D)
Melezitose
MelR1

Figure 4.4 Panel A: Representation of the melezitose utilisation cluster of *B. breve* UCC2003 and DNA fragments used in electrophoretic mobility shift assays (EMSAs) for the *mel*A and *Bbr_1862* promoter regions. Plus and minus signs indicate whether or not MelR1 or MelR2 was able to bind to the *mel*A or *Bbr_1862* DNA fragments, respectively. Panel B: EMSAs showing (i) MelR1 interaction with DNA fragments encompassing fragment M1 and (ii) MelR2 interaction with DNA fragments encompassing fragment K1. Panel C: EMSAs illustrating MelR1 interaction with (i) melAIR and (ii) mutated derivative mu-melAIR and MelR2 interaction with (iii) *Bbr1862*IR and (IV) mutated derivative mu-*Bbr1862*IR. In each panel lane - represents a binding reaction to which no protein was added, while the remaining lanes represent binding reactions with the respective DNA probes incubated with increasing amounts of protein at concentrations ranging from 0.04 nM to 0.01µM. Each successive lane, from right to left, corresponds to a doubling in the concentration of each protein. Panel D: EMSAs showing MelR1 interaction with the DNA fragment M1 with the addition of melezitose at concentrations ranging from 2.5-20 mM.

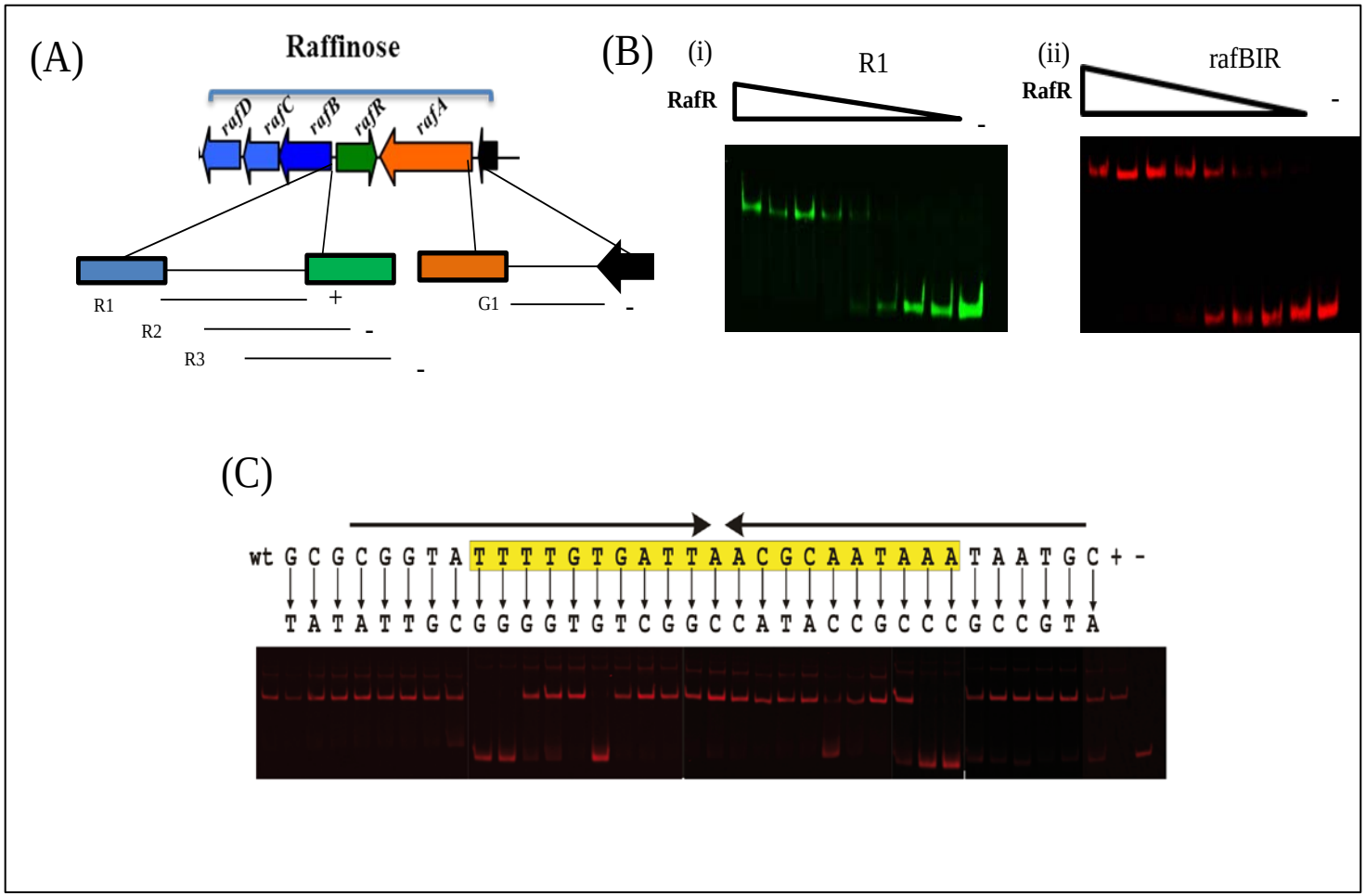
(A)
Raffinose
rafD
rafC
rafB
rafR
rafA
R1
R2
R3
G1
+
-
-
-
(B)
(i)
R1
RafR
-
(ii)
rafBIR
RafR
-
(C)
wt G C G C G G T A T T T T G T G A T T A A C G C A A T A A A T A A T G C + -
T A T A T T G C G G G G T G T C G G C C A T A C C G C C C G C C G T A

Figure 4.5 Panel A: Representation of the raffinose utilisation cluster of *B. breve* UCC2003 and DNA fragments used in electrophoretic mobility shift assays (EMSAs) for the *rafB* and *raf*A promoter regions. Plus and minus signs indicate whether or not RafA was able to bind to the *rafB* or *raf*A DNA fragments, respectively. Panel B: EMSAs showing RafR interaction with DNA fragments encompassing fragment (i) R1 and (ii)the annealed oligonucleotides representing rafBIR. In each panel lane - represents a binding reaction to which no protein was added, while the remaining lanes represent binding reactions with the respective DNA probes incubated with increasing amounts of protein at concentrations ranging from 0.04 nM 0.01 µM. Each successive lane, from right to left, corresponds to a doubling in the concentration of each protein. Panel C: EMSAs showing RafR, interaction with mutated operator motif of the *raf*B promoter region. DNA fragments were obtained by PCR using IRD700 labelled primers. The operator sequence and incorporated mutations are shown above the image. wt and +: original promoter sequence, - : no added protein. In grey the conserved ROK motif. Arrows depict the inverted repeat structure of the motif.

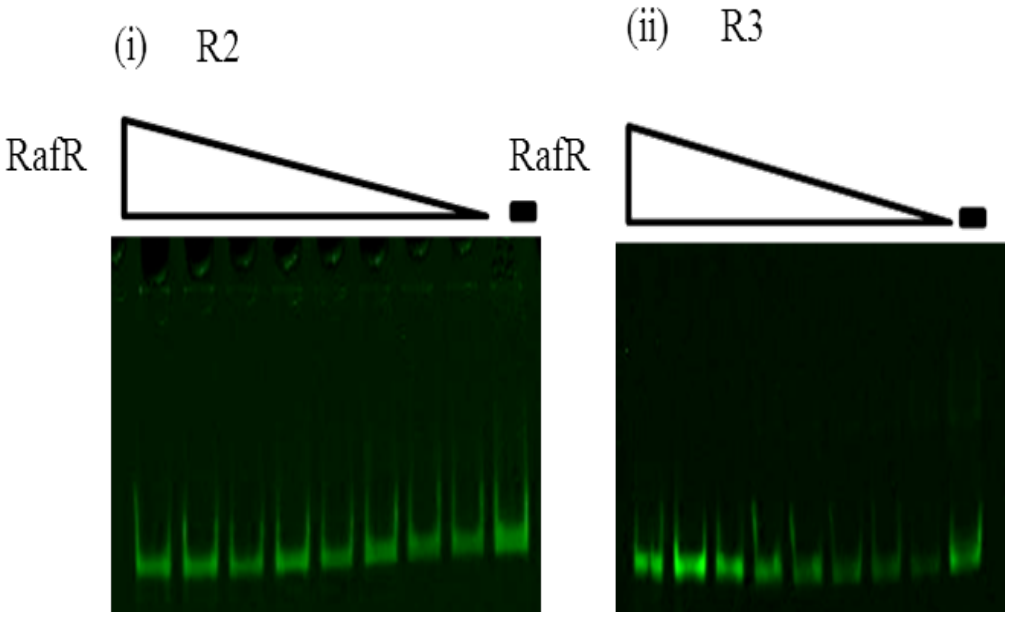

Figure 4.6 EMSAs showing RafR interaction with DNA fragments encompassing fragment R2 (panel i) or R3 (panel ii)**.**

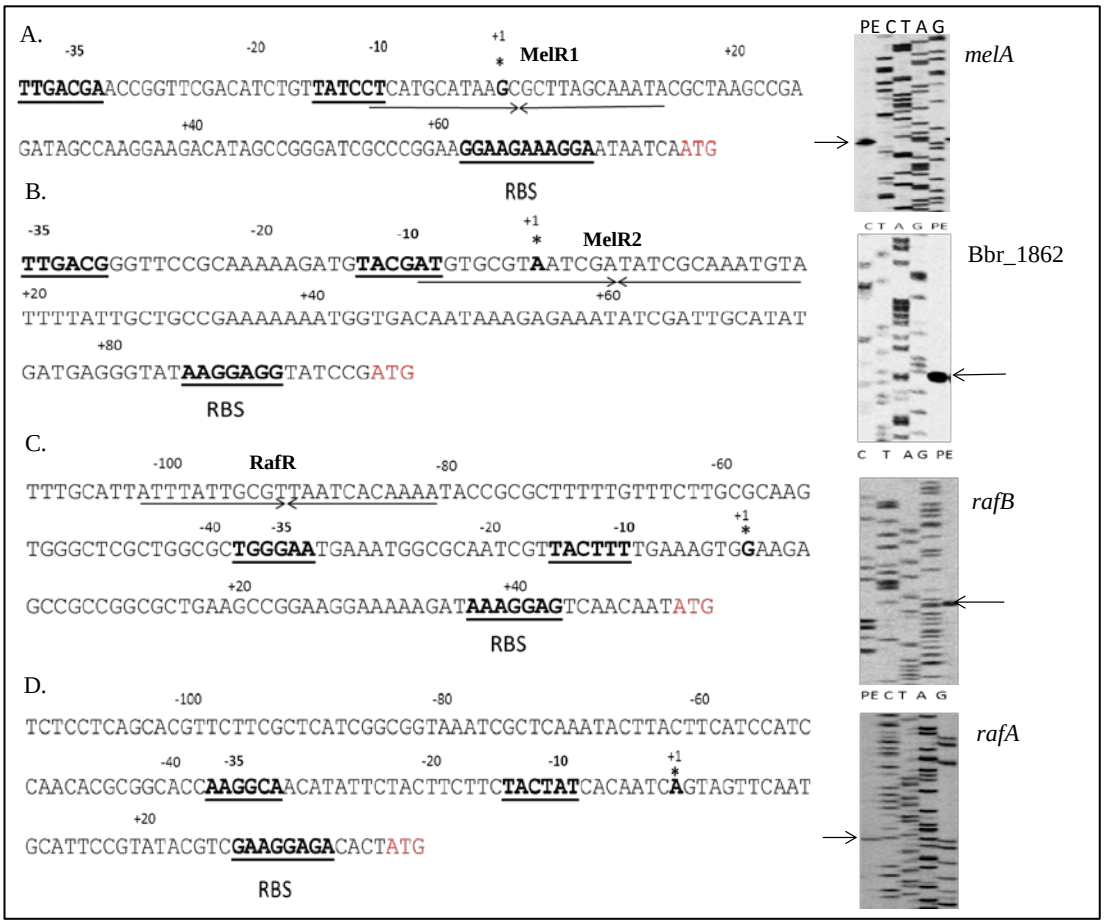

Figure 4.7 Schematic representation of the *melA* (panel A), *Bbr_1862* (panel B), *rafB* (panel C) *rafA* (panel D) promoter regions. Boldface type and underlining indicate the −10 and −35 hexamers as deduced from the primer extension results and ribosomal binding site (RBS); the transcriptional start sites (TSS) are indicated by asterisks. Arrows under sequence in bold indicate the inverted repeat sequence that represents the MelR1, MelR2 or RafR binding sequences as displayed above the arrows.

4.8 SUPPLEMENTARY MATERIAL

Table S4.1 Oligonucleotide primers used in this study

Primer	Sequence[a]	Size bp
melR1fHd3	TGCGGA**AAGCTT**ATTCTTTACACGGTTACC	380
melR1rxba1	CTATGC**TCTAGA**CGATGACCTGTTTGGCCGA	
melR2fHd3	TGCGGA**AAGCTT**ACTTGCTATCATACCGAC	402
melR2rxba1	CTATGC**TCTAGA**CCTGTACTGCGATGATGT	
rafRfHd3	GCTA**AAGCTT**GCAGACCGCATT	530
rafRrxba1	GTCA**TCTAGA**GCACATGCCCCACTAGGC	
melR1-confirm	GCTCGCGAACGGCTGGC	2600
Tetwsal1F	TCAGCT**GTCGAC**ATGCTCATGTACGGTAAG	
melR2-confirm	TCACTGGCTGGCTGGTTC	2600
Tetwsal1F	TCAGCT**GTCGAC**ATGCTCATGTACGGTAAG	
rafR-confirm	GCATAGATCTTCTGACGCCAATGCGATTCC	2600
Tetwsal1F	TCAGCT**GTCGAC**ATGCTCATGTACGGTAAG	
Tetwsal1R	GCGACG**GTCGAC**CATTACCTTCTGAAACATA	
melR1EcorVF	GACAAG**GATATC**ATGCATCACCATCACCATCACCATCACCATCACAAACGCGCGACCATCAAC	1001
melR1Xba1	GACAAG**TCTAGA**CTATTCCCGCTCTGGCCC	
melR2Nco1F	TGCACG**CCATGG**GCCATCACCATCACCATCACCATCACCATCACAGCGAACCAACAATCTATG	1002
melR2Xba1R	CTATGC**TCTAGA**TTAATCTCGAGGTGCTGG	
rokFfbglII	GCAT**AGATCT**TCTGACGCCAATGCGATTCC	
rokRrpst1	CGTA**CTGCAG**CCTACCCAATATGCTTCAC	1206
rafAPEFDNA	CTGAAGTGCATGGCGGCG	

Primer	Sequence[a]	Size bp
rafAPERDNA	GCATCATACGCAGCGAGCAAG	378
rafBPEFDNA	CATTGGCGTCAGACATACTTAAG	
rafBPERDNA	GAGAACCCGCCGACCACG	735
melAPEFDNA	AGCCTTCGAAACGTCATCC	
melAPERDNA	GGTGATGTTCGGGTACT	514
1862PEFDNA	GATTGTTGGTTCGCTCATGGG	
1862PERDNA	CTTTGATGGTGATGTTTG	471

[a] Sequences of restriction enzyme sites are indicated in bold

Table S4.2 IRD800 primers used to generate PCR products and primer extension products in this study

Name	Primer	Product size (bp)
melAIRD800f	AGCCTTCGAAACGTCATCC	
melAIRD800r	CAGACATCAGACCGGCAATAAC	322
melAIRD800SET1F	TCGCGAGAAGGAAGTGC	
melAIRD800SET1R	GTTGGCTTTGTTGCTGGA	312
melAIRD800SET2F	TCGGCCATTTGTTTGCG	
melAIRD800SET2R	GCCAGCGCCGTCCCACA	302
melAIRD800SET3F	TAAGCCGAGATAGCCAAG	
melAIRD800SET3R	GGTGATGTTCGGGTACT	302
1862IRD800irdf	GATTGTTGGTTCGCTCATGGG	
1862IRD800irdr	GAGCCAGGACCACATGCAGT	310
1862IRD800SET1F	GAAGTTAAAGTCGCCATTG	
1862IRD800SET1R	CTAAATCAGTGCTGACATC	300
1862IRD800SET2F	CGTAATCGATATCGCAAATG	
1862IRD800SET2R	GCGCCTCATCTACAGTCTTC	299
1862IRD800SET3F	GATGAGGGTATAAGGAGGTATC	
1862IRD800SET3R	CTTTGATGGTGATGTTTG	268
rafBIRD800irdf	CATTGGCGTCAGACATACTTAAG	
rafBIRD800irdr	GTCTAGGCGTTGAATGTG	222
rafBIRD800SET1F	AAGTGGGCTCGCTGGCG	
rafBIRD800SET1R	CACGCCACAGCAGCAGT	178
rafBIRD800SET2F	GCTGAAGCCGGAAGGAA	
rafBIRD800SET2R	AGCGTAACGGTACCTGC	172
rafAIRD800f	GACTCTCCTCAGCACGTTCTTC	
rafAIRD800r	GCCCTTAAGATCGCCGA	223
rafAPERP1	CTCACCGCCATCTATCTTG	
rafAPERP2	CGACGTATACGGAATGCATTG	
rafBPERP1	CACGCCACAGCAGCAGT	
rafBPERP2	AGCGTAACGGTACCTGC	
melAPERP1	GTTGGCTTTGTTGCTGGA	
melAPERP1	GCCAGCGCCGTCCCACA	
1862PERP1	CTAAATCAGTGCTGACATC	
1862PERP1	GCGCCTCATCTACAGTCTTC	

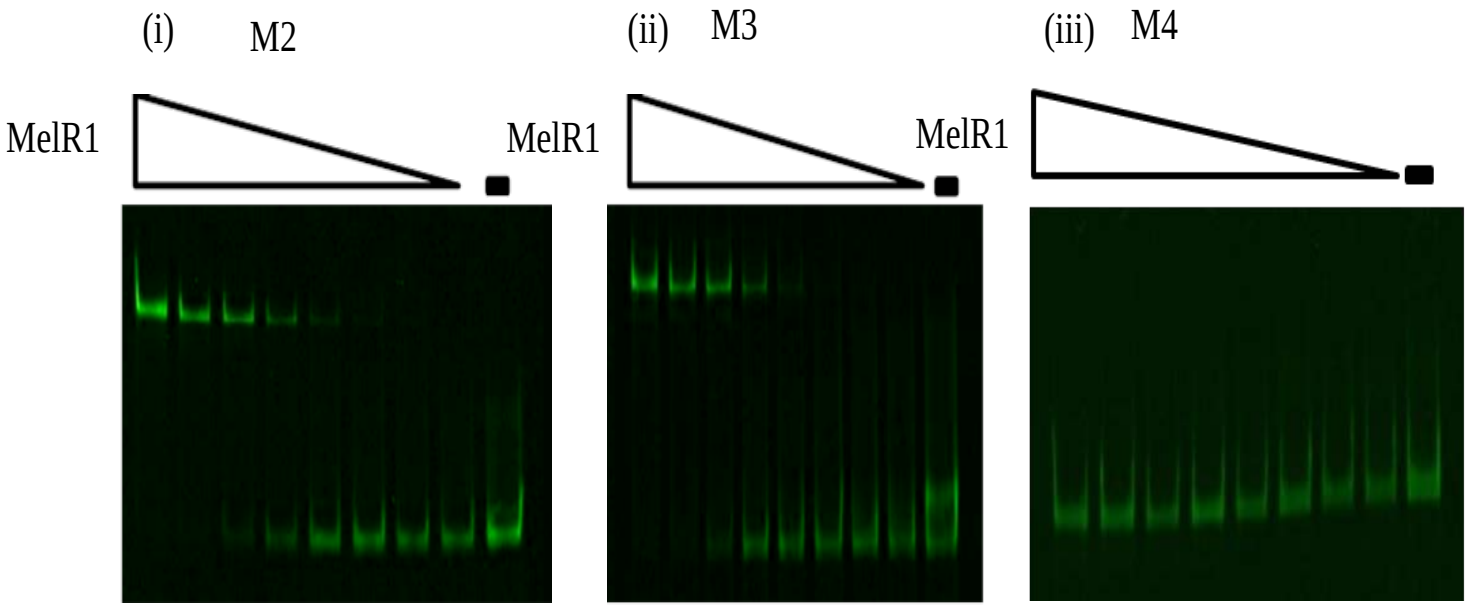

Figure S4.1 EMSAs showing MelR1 interaction with DNA fragments encompassing fragments (i) M2 (ii) M3 (iii) M4.

Table S4.3 IRD800 labelled annealed primers used to generate wild type sequence (wt) and also a series of mutated sequences (m) in relation to *melA* binding sequence.

Wt CTCATGCATAAGCGCTTAGCAAATACGCTAAGCCGAGATAGCCAAG
M1 CTCATGCATAAGCTATTAGCAAATACGCTAAGCCGAGATAGCCAAG
M2 CTCATGCATAAGCGCTTAGCAAATACGAGAAGCCGAGATAGCCAAG
M3 CTCATTAATAAGCGCTTAGCAAATACGCTAAGCCGAGATAGCCAAG
M4 CTCATGCATAAGCGCTTAGCAAATAATCTAAGCCGAGATAGCCAAG

Red lettering indicates the mutated base as compared to the WT sequence. All primers were annealed to reverse complementary sequence, forward primers indicated in table.

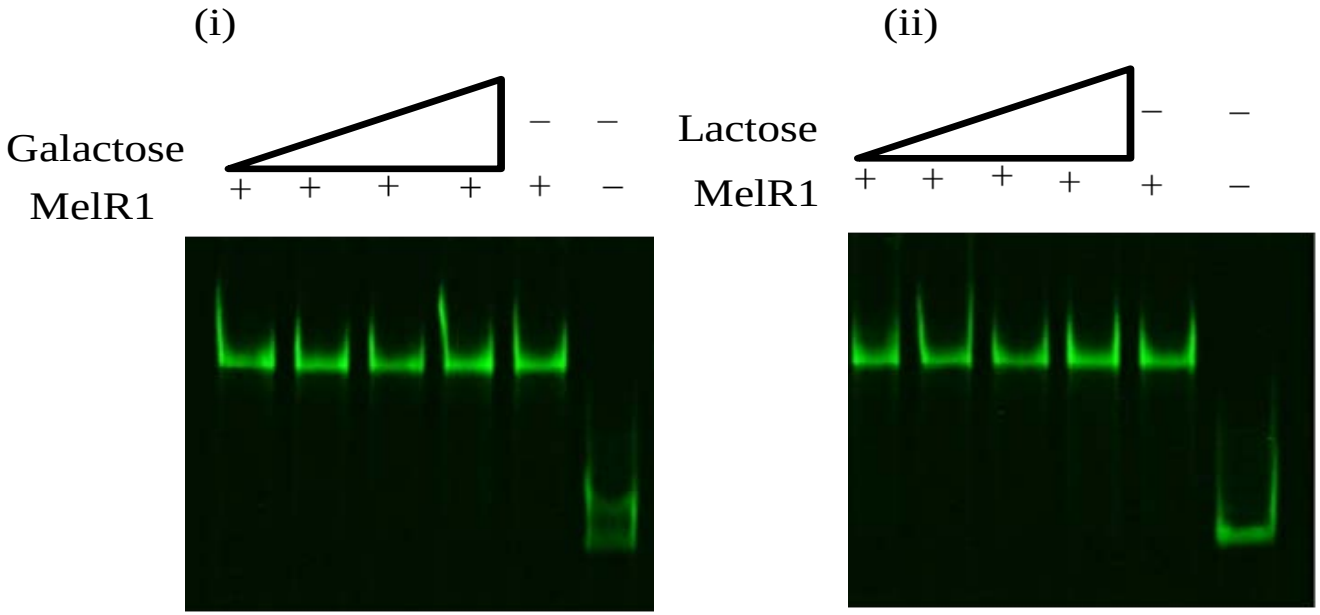

Figure S4.2 EMSAs showing MelR1 interaction with the DNA fragment M1 with the addition of (i) galactose and (ii) lactose at concentrations ranging from 2.5-20 mM.+ carbohydrate plus protein, + - protein no carbohydrate, -- no protein no sugar.

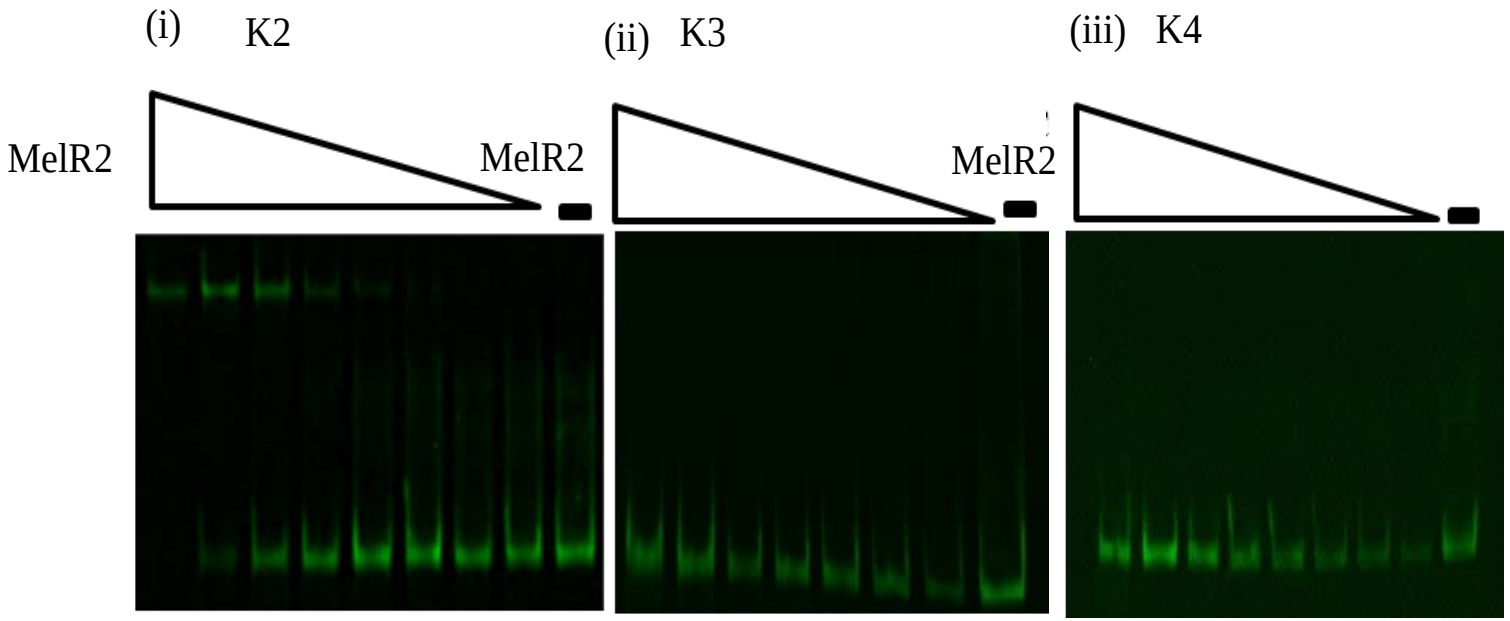

Figure S4.3 EMSAs showing MelR2 interaction with DNA fragments encompassing fragments (i) K2 (ii) K3 and (iii) K4.

Table S4.4 IRD800 labelled annealed primers used to generate wild type sequence (wt) and also a series of mutated sequences (m) in relation to *Bbr_1862* binding sequence.

Wt	TGACGGGTTCCGCAAAAAGATGTACGATGTGCGTAATCGATATCGCAAATG
M1	TGACGGGTTCCGCAAAAAGATGTACGCGGTGCGTAATCGATATCGCAAATG
M2	TGACGGGTTCCTAAAAAAGATGTACGATGTGCGTAATCGATATCGCAAATG
M3	TGACGGGTTCCGCAAAAAGATGTACGATGTGCGTACGCGATATCGCAAATG

Red lettering indicates the mutated base as compared to the WT sequence. All primers were annealed to reverse complementary sequence, forward primers indicated in table.

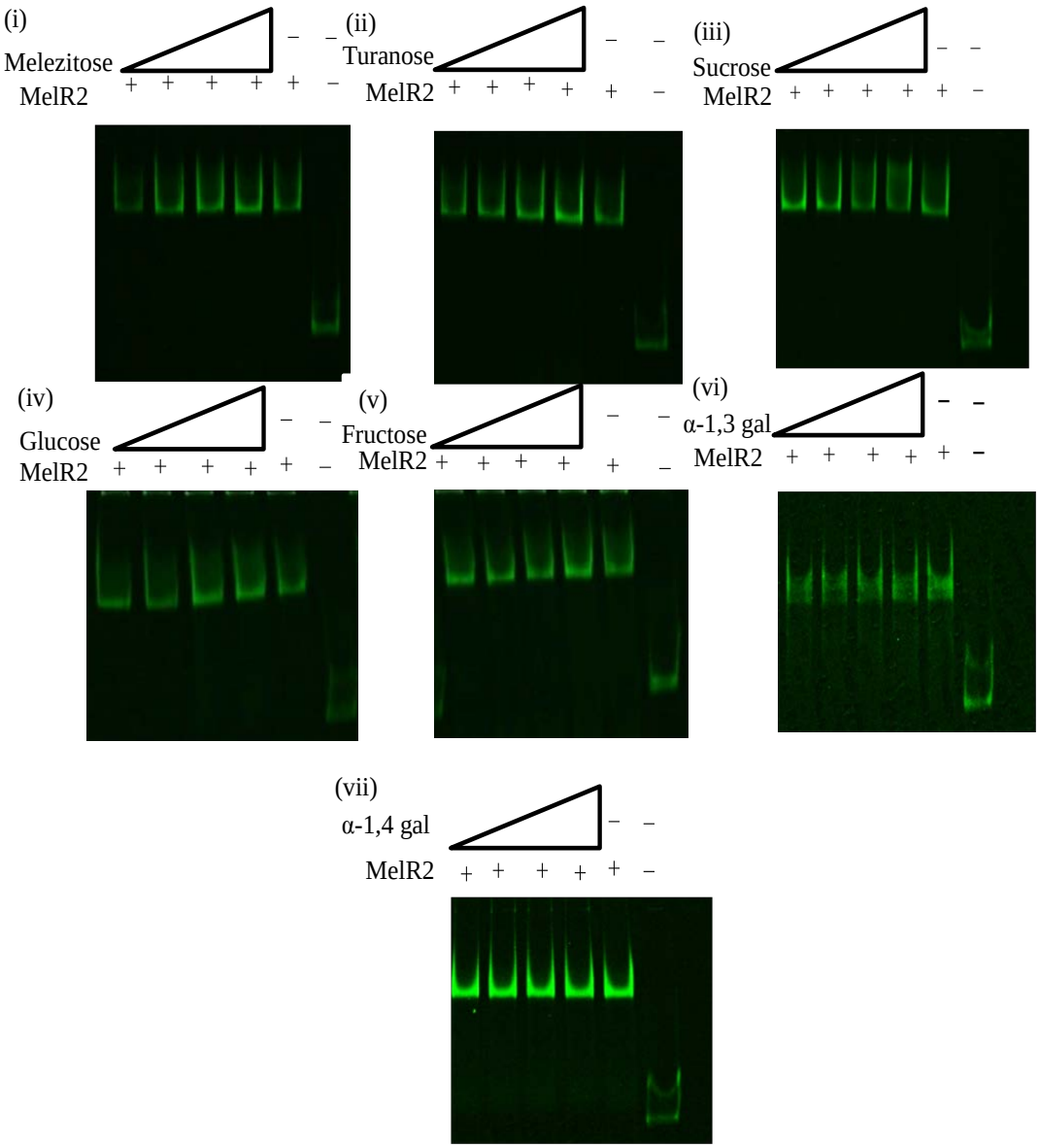

Figure S4.4 EMSAs showing MelR2 interaction with the DNA fragment K1 with the addition of (i) Melezitose (ii) Turanose (iii) Sucrose (iv) Glucose (v) fructose (vi) α-1,3 galactobiose (vii) α-1,4 galactobiose at concentrations ranging from 2.5-20 mM.+ carbohydrate plus protein, + - protein no carbohydrate, -- no protein no sugar.

Table S4.5 IRD700 labelled primers used to generate annealed wild type sequence (wt) in combination with reverse complementary primer, only forward primer indicated, and also a series of mutated sequences (m) by PCR used in combination with RV in relation to RafR binding sequence.

RV	GCGATGCGTATGCTCGGATC
WT	GCGCGGTATTTTGTGATTAACGCAATAAATAATGC
M1	TCGCGGTATTTTGTGATTAACGCAATAAATAATGC
M2	GAGCGGTATTTTGTGATTAACGCAATAAATAATGC
M3	GCTCGGTATTTTGTGATTAACGCAATAAATAATGC
M4	GCGAGGTATTTTGTGATTAACGCAATAAATAATGC
M5	GCGCTGTATTTTGTGATTAACGCAATAAATAATGC
M6	GCGCGTTATTTTGTGATTAACGCAATAAATAATGC
M7	GCGCGGGATTTTGTGATTAACGCAATAAATAATGC
M8	GCGCGGTCTTTTGTGATTAACGCAATAAATAATGC
M9	GCGCGGTAGTTTGTGATTAACGCAATAAATAATGC
M10	GCGCGGTATGTTGTGATTAACGCAATAAATAATGC
M11	GCGCGGTATTGTGTGATTAACGCAATAAATAATGC
M12	GCGCGGTATTTGGTGATTAACGCAATAAATAATGC
M13	GCGCGGTATTTTTTGATTAACGCAATAAATAATGC
M14	GCGCGGTATTTTGGGATTAACGCAATAAATAATGC
M15	GCGCGGTATTTTGTTTATTAACGCAATAAATAATGC
M16	GCGCGGTATTTTGTGCTTAACGCAATAAATAATGC
M17	GCGCGGTATTTTGTGAGTAACGCAATAAATAATGC
M18	GCGCGGTATTTTGTGATGAACGCAATAAATAATGC
M19	GCGCGGTATTTTGTGATTCACGCAATAAATAATGC
M20	GCGCGGTATTTTGTGATTACCGCAATAAATAATGC
M21	GCGCGGTATTTTGTGATTAAAGCAATAAATAATGC
M22	GCGCGGTATTTTGTGATTAACTCAATAAATAATGC
M23	GCGCGGTATTTTGTGATTAACGAAATAAATAATGC
M24	GCGCGGTATTTTGTGATTAACGCCATAAATAATGC
M25	GCGCGGTATTTTGTGATTAACGCACTAAATAATGC
M26	GCGCGGTATTTTGTGATTAACGCAAGAAATAATGC
M27	GCGCGGTATTTTGTGATTAACGCAATCAATAATGC
M28	GCGCGGTATTTTGTGATTAACGCAATACATAATGC
M29	GCGCGGTATTTTGTGATTAACGCAATAACTAATGC
M30	GCGCGGTATTTTGTGATTAACGCAATAAAGAATGC
M31	GCGCGGTATTTTGTGATTAACGCAATAAATCATGC
M32	GCGCGGTATTTTGTGATTAACGCAATAAATACTGC
M33	GCGCGGTATTTTGTGATTAACGCAATAAATAAGGC
M34	GCGCGGTATTTTGTGATTAACGCAATAAATAATTC
M35	GCGCGGTATTTTGTGATTAACGCAATAAATAATGA

Red lettering indicates the mutated base as compared to the WT sequence.

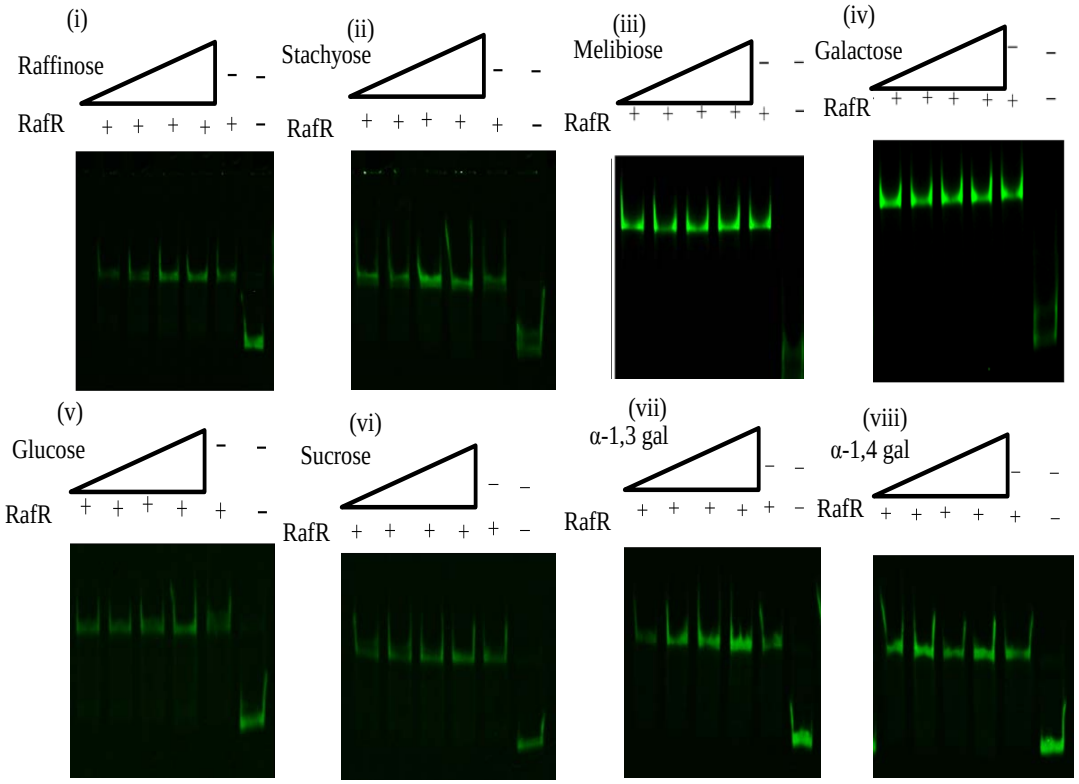

Figure S4.5 EMSAs showing RafR interaction with the DNA fragment R1 with the addition of (i) raffinose (ii) stachyose (iii) melibiose (iv) galactose (v) glucose (vi) sucrose (vii) α-1,3 galactobiose (viii) α-1,4 galactobiose at concentrations ranging from 2.5-20 mM.+ carbohydrate plus protein, + - protein no carbohydrate, -- no protein no sugar.

4.9 REFERENCES

Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. (1990). Basic local alignment search tool. *Journal of Molecular Biology* **215**, 403-410.

Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J., Zhang, Z., Miller, W. & Lipman, D. J. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Research* **25**, 3389-3402.

Alvarez-Martin, P., O'Connell-Motherway, M., van Sinderen, D. & Mayo, B. (2007). Functional analysis of the pBC1 replicon from *Bifidobacterium catenulatum* L48. *Applied Microbiology and Biotechnology* **76**, 1395-1402.

Bacon, J. S. & Dickinson, B. (1957). The origin of melezitose: a biochemical relationship between the lime tree (*Tilia spp.*) and an aphis (*Eucallipterus tiliae L.*). *The Biochemical Journal* **66**, 289-297.

Bouhnik, Y., Vahedi, K., Achour, L., Attar, A., Salfati, J., Pochart, P., Marteau, P., Flourié, B., Bornet, F. & Rambaud, J.C. (1999). Short-chain fructo-oligosaccharide administration dose-dependently increases fecal bifidobacteria in healthy humans. *The Journal of Nutrition* **129**, 113-116.

Bouhnik, Y., Attar, A., Joly, F. A., Riottot, M., Dyard, F. & Flourie, B. (2004). Lactulose ingestion increases faecal bifidobacterial counts: A randomised double-blind study in healthy humans. *European Journal of Clinical Nutrition* **58**, 462-466.

Cecchini, D. A., Laville, E., Laguerre, S. Robe, P., Leclerc, M., Doré, J., Henrissat, B., Remaud-Siméon, M., Monsan, P. & Potocki-Véronèse, G. (2013). Functional Metagenomics Reveals Novel Pathways of Prebiotic Breakdown by Human Gut Bacteria. *PLoS ONE* **8**, e72766.

Collado-Vides, J., Magasanik, B. & Gralla, J. D. (1991). Control site location and transcriptional regulation in *Escherichia coli*. *Microbiological Reviews* **55**, 371-394.

Costabile, A., Klinder, A., Fava, F., Napolitano, A., Fogliano, V., Leonard, C., Gibson, G. R. & Tuohy, K. M. (2008). Whole-grain wheat breakfast cereal has a prebiotic effect on the human gut microbiota: a double-blind, placebo-controlled, crossover study. *The British Journal of Nutrition* **99**, 110-120.

De Man, J. C., Rogosa, M. & Sharpe, M. E. (1960). A medium for the cultivation of *lactobacilli. Journal of Applied Microbiology* **23**, 130-135.

de Ruyter, P. G., Kuipers, O. P., Beerthuyzen, M. M., van Alen-Boerrigter, I. & de Vos, W. M. (1996). Functional analysis of promoters in the nisin gene cluster of *Lactococcus lactis*. *Journal of Bacteriology* **178**, 3434-3439.

Deutscher, J. (2008). The mechanisms of carbon catabolite repression in bacteria. *Current Opinion in Microbiology* **11**, 87-93.

Erdogan, O., Tanyeri, B., Torun, E., Gonullu, E., Arslan, H.,Erenberk, U. & Oktem, F. (2012). The Comparition of the Efficacy of Two Different Probiotics in Rotavirus Gastroenteritis in Children. *Journal of Tropical Medicine* **2012**, 5.

Fukami-Kobayashi, K., Tateno, Y. & Nishikawa, K. (2003). Parallel evolution of ligand specificity between LacI/GalR family repressors and periplasmic sugar-binding proteins. *Molecular Biology and Evolution* **20**, 267-277.

Garcia De La Nava, J., Santaella, D. F., Alba, J. C., Carazo, J. M., Trelles, O. & Pascual-Montano, A. (2003). Engene: the processing and exploratory analysis of gene expression data. *Bioinformatics* **19**, 657-658.

Gibson, G. R. & Roberfroid, M. B. (1995). Dietary modulation of the human colonic microbiota: introducing the concept of prebiotics. *The Journal of Nutrition* **125**, 1401-1412.

Gough, E., Shaikh, H. & Manges, A. R. (2011). Systematic Review of Intestinal Microbiota Transplantation (Fecal Bacteriotherapy) for Recurrent *Clostridium difficile* Infection. *Clinical Infectious Diseases* **53**, 994-1002.

Tissier, H. (1906). The treatment of intestinal infections by the method of transformation of the bacterial intestinal flora. *CR Soc Biol* **60**, 359-361.

Hellman, L. M. & Fried, M. G. (2007). Electrophoretic mobility shift assay (EMSA) for detecting protein-nucleic acid interactions. *Nature Protocols* **2**, 1849-1861.

Irwin, N. & Ptashne, M. (1987). Mutants of the Catabolite Activator Protein of *Escherichia coli* that are Specifically Deficient in the Gene-Activation Function. *Proceedings of the National Academy of Sciences of the United States of America* **84**, 8315-8319.

Kuipers, O. P., Beerthuyzen, M. M., Siezen, R. J. & de Vos, W. M. (1993). Characterisation of the nisin gene cluster nisABTCIPR of *Lactococcus lactis* . Requirement of expression of the nisA and nisI genes for development of immunity. *European Journal of Biochemistry* **216**, 281-291.

Laemmli, U. K. (1970). Cleavage of Structural Proteins during the Assembly of the Head of Bacteriophage T4. *Nature* **227**, 680-685.

Law, J., Buist, G., Haandrikman, A., Kok, J., Venema, G. & Leenhouts, K. (1995). A system to generate chromosomal mutations in *Lactococcus lactis* which allows fast analysis of targeted genes. *Journal of Bacteriology* **177**, 7011-7018.

Long, A. D., Mangalam, H. J., Chan, B. Y., Tolleri, L., Hatfield, G. W. & Baldi, P. (2001). Improved statistical inference from DNA microarray data using analysis of variance and a Bayesian statistical framework. Analysis of global gene expression in *Escherichia coli* K12. *Journal of Biological Chemistry* **276**, 19937-19944.

Macfarlane, G. T., Steed, H. & Macfarlane, S. (2008). Bacterial metabolism and health-related effects of galacto-oligosaccharides and other prebiotics. *Journal of Applied Microbiology* **104**, 305-344.

Madan Babu, M. & Teichmann, S. A. (2003). Functional determinants of transcription factors in *Escherichia coli*: protein families and binding sites. *Trends in Genetics* **19**, 75-79.

Maze, A., O'Connell-Motherway, M., Fitzgerald, G. F., Deutscher, J. & van Sinderen, D. (2007). Identification and characterisation of a fructose phosphotransferase system in *Bifidobacterium breve* UCC2003. *Applied and Environmental Microbiology* **73**, 545-553.

Mierau, I., Leij, P., van Swam, I., Blommestein, B., Floris, E., Mond, J. & Smid, E. (2005). Industrial-scale production and purification of a heterologous protein in *Lactococcus lactis* using the nisin-controlled gene expression system NICE: The case of lysostaphin. *Microbial Cell Factories* **4**, 15.

Miyake, T., Watanabe, K., Watanabe, T. & Oyaizu, H. (1998). Phylogenetic analysis of the genus *Bifidobacterium* and related genera based on 16S rDNA sequences. *Microbiology and Immunology* **42**, 661-667.

O'Connell, K. J., O'Connell Motherway, M., O'Callaghan, J., Fitzgerald, G. F., Ross, R. P., Ventura, M., Stanton, C. & van Sinderen, D. (2013). Metabolism of four α-glycosidic linkage-containing oligosaccharides by *Bifidobacterium breve* UCC2003. *Applied and Environmental Microbiology*. **79,** 6280-6292.

O'Connell Motherway, M., Fitzgerald, G. F., Neirynck, S., Ryan, S., Steidler, L. & van Sinderen, D. (2008). Characterisation of ApuB, an extracellular type II amylopullulanase from *Bifidobacterium breve* UCC2003. *Applied and Environmental Microbiology* **74**, 6271-6279.

O'Connell Motherway, M., O'Driscoll, J., Fitzgerald, G. F. & Van Sinderen, D. (2009). Overcoming the restriction barrier to plasmid transformation and targeted mutagenesis in *Bifidobacterium breve* UCC2003. *Microbial Biotechnology* **2**, 321-332.

O'Connell Motherway, M., Fitzgerald, G. F. & van Sinderen, D. (2011a). Metabolism of a plant derived galactose-containing polysaccharide by *Bifidobacterium breve* UCC2003. *Microbial Biotechnology* **4**, 403-416.

O'Connell Motherway, M., Zomer, A., Leahy, S. C. O'Connell Motherway, M., Zomer, A., Leahy, S. C., Reunanen, J., Bottacini, F., Claesson, M.J., O'Brien, F., Flynn, K., Casey, P.G., Munoz, J.A., Kearney, B., Houston, A.M., O' Mahony, C., Higgins, DG., Shanahan, F., Palva, A., de Vos, W.M., Fitzgerald, G.F., Ventura, M., O'Toole, P.W. & van Sinderen D. (2011b). Functional genome analysis of *Bifidobacterium breve* UCC2003 reveals type IVb tight adherence (Tad) pili as an essential and conserved host-colonisation factor. *Proceedings of the National Academy of Sciences of the United States of America* **108**, 11217-11222.

O'Riordan, K. & Fitzgerald, G. F. (1999). Molecular characterisation of a 5.75-kb cryptic plasmid from *Bifidobacterium breve* NCFB 2258 and determination of mode of replication. *FEMS Microbiology Letters* **174**, 285-294.

Parche, S., Beleut, M., Rezzonico, E., Jacobs, D., Arigoni, F., Titgemeyer, F. & Jankovic, I. (2006). Lactose-over-glucose preference in *Bifidobacterium longum* NCC2705: glcP, encoding a glucose transporter, is subject to lactose repression. *Journal of Bacteriology* **188**, 1260-1265.

Parker, M. W., Schaffzin, J. K., Lo Vecchio, A. Yau, C., Vonderhaar, K., Guiot, A., Brinkman, WB., White, CM., Simmons, J.M,, Gerhardt, W.E., Kotagal, U.R. & Conway, P.H. (2013). Rapid adoption of *Lactobacillus rhamnosus* GG for acute gastroenteritis. *Pediatrics* **131**, S96-102.

Plumbridge, J. (2001). Regulation of PTS gene expression by the homologous transcriptional regulators, Mlc and NagC, in *Escherichia coli* (or how two similar repressors can behave differently). *Journal of Molecular Microbiology and Biotechnology* **3**, 371-380.

Pokusaeva, K., Neves, A. R., Zomer, A., O'Connell-Motherway, M., MacSharry, J., Curley, P., Fitzgerald, G. F. & van Sinderen, D. (2010). Ribose utilisation by the human commensal *Bifidobacterium breve* UCC2003. *Microbial Biotechnology* **3**, 311-323.

Pokusaeva, K., O'Connell-Motherway, M., Zomer, A., Macsharry, J., Fitzgerald, G. F. & van Sinderen, D. (2011). Cellodextrin utilisation by *Bifidobacterium breve* UCC2003. *Applied and Environmental Microbiology* **77**, 1681-1690.

Rohlke, F. & Stollman, N. (2012). Fecal microbiota transplantation in relapsing Clostridium difficile infection. *Therapeutic Advances in Gastroenterology* **5**, 403-420.

Ruckert, C., Milse, J., Albersmeier, A., Koch, D., Puhler, A. & Kalinowski, J. (2008). The dual transcriptional regulator CysR in *Corynebacterium glutamicum* ATCC 13032 controls a subset of genes of the McbR regulon in response to the availability of sulphide acceptor molecules. *BMC Genomics* **9**, 483.

Rutherford, K., Parkhill, J., Crook, J., Horsnell, T., Rice, P., Rajandream, M. A. & Barrell, B. (2000). Artemis: sequence visualisation and annotation. *Bioinformatics* **16**, 944-945.

Ryan, S. M., Fitzgerald, G. F. & van Sinderen, D. (2005). Transcriptional Regulation and Characterisation of a Novel β-Fructofuranosidase-Encoding Gene from *Bifidobacterium breve* UCC2003. *Applied and Environmental Microbiology* **71**, 3475-3482.

Sambrook, J., Fritsch, E. F. & Maniatis, T. (1989). *Molecular cloning: a laboratory manual*: Cold Spring Harbor Laboratory.

Scardovi, V. & Trovatelli, I. D. (1965). The fructose-6-phosphate shunt as a peculiar pattern of hexose degradation in the genus *Bifidobacterium*. *Annals of Microbiology* **15**, 19-29.

Swint-Kruse, L. & Matthews, K. S. (2009). Allostery in the LacI/GalR family: variations on a theme. *Current Opinion in Microbiology* **12**, 129-137.

Terzaghi, B. E. & Sandine, W. E. (1975). Improved medium for lactic *Streptococci* and their bacteriophages. *Applied Microbiology* **29**, 807-813.

Tissier, H. (1900). Recherchers sur la flora intestinale normale pathologique du nourisson. *Thesis, University of Paris, Paris France*.

Titgemeyer, F., Reizer, J., Reizer, A. & Saier, M. H. (1994). Evolutionary relationships between sugar kinases and transcriptional repressors in bacteria. *Microbiology* **140**, 2349-2354.

van Hijum, S. A. F. T., Garcia De La Nava, J., Trelles, O., Kok, J. & Kuipers, O. P. (2003). MicroPreP: a cDNA microarray data pre-processing framework. *Applied Bioinformatics* **2**, 241-244.

van Hijum, S. A. F. T., De, J. A., Baerends, R. J., Karsens, HA., Kramer, NE., Larsen, R., den Hengst, C.D., Albers, C.J., Kok, J. & Kuipers, O.P. (2005). A generally applicable validation scheme for the assessment of factors involved in reproducibility and quality of DNA-microarray data. *BMC Genomics* **6**, 77.

Ventura, M., Zink, R., Fitzgerald, G. F. & van Sinderen, D. (2005). Gene Structure and Transcriptional Organisation of the dnaK Operon of *Bifidobacterium breve* UCC 2003 and Application of the Operon in Bifidobacterial Tracing. *Applied and Environmental Microbiology* **71**, 487-500.

Ventura, M., Canchaya, C., Tauch, A., Chandra, G., Fitzgerald, G. F., Chater, K. F. & van Sinderen, D. (2007). Genomics of *Actinobacteria*: Tracing the Evolutionary History of an Ancient Phylum. *Microbiology and Molecular Biology Reviews* **71**, 495-548.

Weickert, M. J. & Adhya, S. (1992). A family of bacterial regulators homologous to Gal and Lac repressors. *The Journal of Biological Chemistry* **267**, 15869-15874.

Wells, J. M., Wilson, P. W. & Le Page, R. W. (1993). Improved cloning vectors and transformation procedure for *Lactococcus lactis*. *The Journal of Applied Bacteriology* **74**, 629-636.

Wilson, C. J., Zhan, H., Swint-Kruse, L. & Matthews, K. S. (2007). The lactose repressor system: paradigms for regulation, allosteric behavior and protein folding. *Cellular and Molecular Life Sciences* **64**, 3-16.

Zhu, Q., Gao, R., Wu, W. & Qin, H. (2013). The role of gut microbiota in the pathogenesis of colorectal cancer. *Tumour biology: the journal of the International Society for Oncodevelopmental Biology and Medicine* **34**, 1285-1300.

Zomer, A., Fernandez, M., Kearney, B., Fitzgerald, G. F., Ventura, M. & van Sinderen, D. (2009). An Interactive Regulatory Network Controls Stress Response in *Bifidobacterium breve* UCC2003. *Journal of Bacteriology* **191**, 7039-7049.

Chapter V

Identification and characterisation of putative replication functions of the megaplasmid pBb7017 from *Bifidobacterium breve* JCM 7017

Pulsed field gel electrophoresis was performed by Dr. Mary O' Connell-Motherway.

Schematic representation of pBb7017 was provided by Francesca Bottacini

Schematic representation of neighbour joining tree provided by Amy O' Callaghan

5.1 ABSTRACT

A nursling stool isolate, *Bifidobacterium breve* JCM 7017 was found to harbour a 190.34 kb megaplasmid, designated pBb7017, the first megaplasmid to be positively identified in bifidobacteria. Characterisation of bifidobacterial probiotic determinants has been hindered by the lack of effective molecular tools, including stable and effective cloning and shuttle vectors. Therefore, in this study we focused on the characterisation of the replication functions of the megaplasmid with the intention of creating an *Escherichia coli/Bifidobacterium* shuttle vector based on the megaplasmid replication functions. The megaplasmid was found to harbour two putative replication or rep_{7017} genes, designated $repA_{7017}$ and $repB_{7017}$. One of these, $repA_{7017}$, encoding a putative 39.3 kDa replication protein $RepA_{7017}$, was subject to an in depth characterisation. Several unsuccessful attempts were made to create an *E. coli*/bifidobacterial shuttle vector based on $repA_{7017}$-based replication functions.

5.2 INTRODUCTION

Members of the genus *Bifidobacterium*, known for their typical bifurcated (bifid) morphology, represent saccharolytic, non-motile, non-sporulating, Gram-positive, anaerobic bacteria, which possess a high GC genome content, and belong to the *Actinobacteria* phylum and *Bifidobacteriaceae* family. Bifidobacteria are commonly found as anaerobic commensals of the gastrointestinal tract (GIT) (Miyake *et al.*, 1998; Ventura *et al.*, 2007). Various strains and species of the genus are considered probiotic bacteria, exhibiting one or more beneficial traits such as eliciting a protective effect against Gram negative pathogens (Fanning *et al.*, 2012; Fukuda *et al.*, 2011), alleviating lactose intolerance (Almeida *et al.*, 2012), prevention of diarrhoea (H.Tissier, 1906), and reducing serum cholesterol (Bordoni *et al.*, 2013; Dambekodi & Gilliland, 1998). Bifidobacteria may also be used as tumour-targeting vectors (Cronin *et al.*, 2010; Cronin *et al.*, 2012), or as a live vaccine (Ma *et al.*, 2012).

Until recently, genetic tools for bifidobacteria had been scarce although in recent times a number of tools, for example those required for cloning, gene disruption and transposon mutagenesis, have been developed (O'Connell Motherway *et al.*, 2009; Ruiz *et al.*, 2013; Sun *et al.*, 2012). Also various *E. coli/Bifidobacterium* shuttle vectors have been constructed to date (Klijn *et al.*, 2006; Missich *et al.*, 1994; Rossi *et al.*, 1996; Rossi *et al.*, 1998; Sangrador-Vegas *et al.*, 2007; Shkoporov *et al.*, 2008). For example, the pBC1 replicon from *Bifidobacterium catenulatum* proved to be functional in several *Bifidobacterium* species, including *Bifidobacterium animalis, Bifidobacterium longum,* and *Bifidobacterium pseudocatenulatum* (Alvarez-Martin *et al.*, 2007), while the pPKCM7 replication functions, originally from *Bifidobacterium asteroides*, were shown to be operating in *B. animalis* subsp.

lactis, *B. longum* NCIMB 8809, *Bifidobacterium pseudolongum* NCIMB 2244, *Bifidobacterium pseudolongum* subsp. *globosum* JCM 5820, *B. pseudocatenulatum* LMG 10505 and *Bifidobacterium dentium* NCFB 2843 (Cronin *et al.*, 2007).

The presence of plasmids in bifidobacterial strains was originally reported over thirty years ago (Sgorbati *et al.*, 1982), although only a minority of characterised bifidobacteria (approximately 20 %) appear to contain plasmids (Sgorbati *et al.*, 1982). A total of 36 plasmids have been completely sequenced to date in *Bifidobacterium*, including 24 plasmids from *B. longum*, three from *B. breve*, two from *Bifidobacterium bifidum*, *B. asteroides* and *Bifidobacterium kashiwanohense*, and just one from *B. pseudolongum* subsp. *globosum*, *B. catenulatum* and *B. pseudocatenulatum* (For a full list of plasmid names and references see page 8 in the Introduction of this thesis).

Characterised bifidobacterial plasmids are relatively small, ranging in size from 1.5 kb to 15 kb (Park *et al.*, 1997; Sgorbati *et al.*, 1982). The majority of the plasmids identified in members of the genus *Bifidobacterium* are predicted to replicate via the so-called rolling circle replication mechanism, although plasmids encoding proteins that are similar to known or predicted theta-type replication proteins have also been identified (Alvarez-Martin *et al.*, 2007; Corneau *et al.*, 2004; Lee & O'Sullivan, 2006; Moon *et al.*, 2009; Takahata *et al.*, 2013).

Commensal lactobacilli have been shown to harbour so-called megaplasmids, i.e. extra-chromosomal genetic elements with a size of 100 kb and larger, one of which, pMP118 from *Lactobacillus salivarius* UCC118, was shown to harbour genes that specify proteins involved in contingency amino acid metabolism, carbohydrate utilisation, Abp118 bacteriocin production and immunity, as well as genes encoding

a bile salt hydrolase and a putative conjugation locus (Claesson *et al.*, 2006; Li *et al.*, 2007; O' Shea *et al.*, 2011). The existence of megaplasmids has previously been suggested for porcine isolates of *B. pseudolongum* subsp. *globosum*, though never conclusively proven (Simpson *et al.*, 2003). Here we report on the replication functions of the 190.34 kb megaplasmid pBb7017 of *B. breve* JCM 7017.

5.3 MATERIALS AND METHODS

Bacterial strains, plasmids and culture conditions

Bacterial strains and plasmids used in this study are detailed in Table 5.1. Bifidobacterial strains were routinely cultured in either de Mann, Rogosa and Sharpe medium (MRS; Difco™, BD, Le Pont de Claix, France) supplemented with 0.05 % cysteine-HCl, or reinforced clostridial medium (RCM; Oxoid Ltd.). Bifidobacterial cultures were incubated at 37^{o}C under anaerobic conditions which were maintained using an Anaerocult oxygen depleting system (Merck, Darmstadt, Germany) in an anaerobic chamber. *Lactococcus lactis* strains were cultivated in M17 broth containing 0.5 % glucose (Terzaghi & Sandine, 1975) at 30^{o}C. *Escherichia coli* strains were cultured in Luria-Bertani broth (LB) (Sambrook *et al.*, 1989) at 37^{o}C with agitation. Where appropriate, growth media contained chloramphenicol (Cm; 5 µg ml^{-1} for *L. lactis*, 10 µg ml^{-1} for *E. coli*, and 2.5 µg ml^{-1} for *B. breve*), erythromycin, tetracycline (Tet; 10 µg ml^{-1} for *E. coli* or *B. breve*) or kanamycin (Km; 50 µg ml^{-1} for *E. coli*).

Nucleotide sequence analysis

Sequence data was obtained from the Artemis-mediated (Rutherford *et al.*, 2000) genome annotations of *B. breve* JCM 7017 (Bottacini *et al.*, unpublished data). Database searches were performed using non-redundant sequences accessible at the National Centre for Biotechnology Information (http://www.ncbi.nlm.nih.gov) using BLASTP or BLASTN (Altschul *et al.*, 1990; Altschul *et al.*, 1997). Sequence verification and analysis were performed using the Seqman and Seqbuilder programs of the DNASTAR software package (DNASTAR, Madison, WI, USA v10.1.2).

Phylogenetic analysis of the $repA_{7017}$ and $repB_{7017}$ genes.

A BLAST alignment of the encoded protein products of the $repA_{7017}$ and $repB_{7017}$ genes of the pBb7017 megaplasmid was performed against the NCBI non redundant database; the first 200 significant hits (e-value < 0.0001) were retrieved and aligned, using Muscle v.3.3 and the relative phylogenetic neighbour-joining tree was computed in Mega5 with a bootstrap analysis of 100 re-samplings.

Pulsed field Gel Electrophoresis (PFGE)

PFGE plug preparation, S1 nuclease treatment, and PFGE were performed as described previously (Li *et al.*, 2007). It is believed that S1 nuclease treatment first introduces a single strand break into a covalently closed plasmid, thereby converting superhelical molecules into an open circle form, after which a second break is introduced at the same site in the complementary strand to produce the linear form of the plasmid (Bai *et al.*, 2003; Beard *et al.*, 1973; Germond *et al.*, 1974),

DNA manipulations

Genomic DNA was isolated from bifidobacteria as previously described (O'Riordan & Fitzgerald, 1999). Minipreparation of plasmid DNA from *E. coli*, *B. breve* or *L. lactis* was achieved using the Qiaprep spin plasmid miniprep kit (Qiagen GmBH, Hilden, Germany). For *B. breve* or *L. lactis* an initial lysis step was incorporated into the plasmid isolation procedure by re-suspending cells in lysis buffer supplemented with lysozyme (30 mg ml^{-1}) followed by incubation at 37^{o}C for 30 min (Sambrook *et al.*, 1989). Procedures for DNA manipulations were performed essentially as described previously (Sambrook *et al.*, 1989). Restriction enzymes and T4 DNA ligase were used according to the supplier's instructions (Roche Diagnostics, Bell

Lane, East Sussex, UK). Synthetic single stranded oligonucleotide primers used in this study are detailed in supplementary Table S5.1, were synthesised by Eurofins (Ebersberg, Germany). Standard PCRs were performed using TaqPCR mastermix (Qiagen), while *B. breve* colony PCR analysis was performed as described previously (O'Connell Motherway *et al.*, 2009). PCR fragments were purified using the Qiagen PCR purification kit (Qiagen). Electroporation of plasmid DNA into *E. coli* was performed as previously described (Sambrook *et al.*, 1989). Electrotransformation of *B. breve* UCC2003 (Maze *et al.*, 2007) and *L. lactis* (Wells *et al.*, 1993) was performed according to published protocols. The correct orientation and integrity of all constructs was verified by DNA sequencing, performed by Eurofins (Ebersberg, Germany).

Plasmid constructions

For heterologous expression of one of the predicted replication proteins of pBb7017, the gene *pl7017_0124* (designated here as $repA_{7017}$), was amplified by PCR employing genomic DNA of *B. breve* JCM 7017 as a template, and using Taq DNA polymerase and primer combination repANco1F and repAXba1R (Supplementary Table S5.1). NcoI and XbaI restriction sites were incorporated at the 5' end of each forward and reverse primer combination, respectively, for the $repA_{7017}$-encompassing primers (Supplementary Table S5.1). In addition, an in frame His_{10}-encoding sequence was incorporated into each of the forward primers to facilitate downstream protein purification using the Ni-NTA affinity system (Qiagen). The $repA_{7017}$-encompassing amplicon was digested with NcoI and XbaI, and ligated into the NcoI and XbaI-digested, nisin-inducible plasmid pNZ8048 (Mierau *et al.*, 2005). The ligation mixture was introduced into *L. lactis* NZ9000 (Table 5.1) by electrotransformation, and transformants were then selected based on

chloramphenicol resistance. The plasmid content of a number of Cm^r transformants was screened by restriction analysis and the integrity of positively identified clones was verified by sequencing, resulting in the construct pNZ-RepA-His**.**

In order to clone the *repA*$_{7017}$ promoter region (termed here as *repAp*), this region which included 321bp upstream of *repA*$_{7017}$ (promoter region) and 218bp of *repA*$_{7017}$ (total 539 bp) was cloned in an *E. coli*/*Bifidobacterium* shuttle vector pAM5 (for primer extension analysis of this promoter), the relevant fragment was amplified by PCR employing genomic DNA of *B. breve* JCM 7017 as a template, and using Taq DNA polymerase and primer combination repApXba1F and repApSph1R (Supplementary Table S5.1). SphI and XbaI restriction sites were incorporated at the 5' end of each forward and reverse primer combination, respectively (Supplementary Table S5.1). The generated 539 bp *repAp*-encompassing amplicon was digested with SphI and XbaI, and ligated into SphI and XbaI-digested pAM5 (Alvarez-Martin *et al.*, 2007). The ligation mixture was introduced into *E. coli* EC101 (Table 5.1) by electrotransformation, and transformants were then selected based on tetracycline resistance. The plasmid content of a number of Tet^r transformants was screened by restriction analysis and the integrity of positively identified clones was verified by sequencing, resulting in plasmid pAM5-repAp**.**

DNA fragments containing *repA*$_{7017}$, *repA*$_{7017}$ plus *pl7017_0123* (designated here as *repA*$_{7017+1}$), and *repA7017* plus *pl7017_0123* plus *pl7017_0122* (designated here as *repA*$_{7017+2}$), all including the *repA*$_{7017}$-upstream, and in the case of *repA*$_{7017+2}$ also some downstream region of the cloned genes, were generated by PCR amplification (Fig. 5.1). This was achieved employing chromosomal DNA of *B. breve* JCM 7017 as a template, and using Taq DNA polymerase and primer combinations 124fsph1-rep,124rsph1-rep, 123+124fsph1-rep and 123+124rsph1-rep, and 122-124fsph1-rep

and 22-124rsph1-rep, respectively, in which SphI restriction sites had been incorporated at the 5' end of each primer (Supplementary table S5.1). The generated amplicons were digested with SphI, and ligated into SphI-digested pBif (Table 5.1). The individual ligation mixtures were introduced into *E. coli* Xl1 blue (Table 5.1) by electrotransformation, and transformants were then selected based on chloramphenicol and kanamycin resistance. The plasmid content of a number of Cm^r and Kn^r transformants was screened by restriction analysis and the integrity of positively identified clones was verified by sequencing. Resulting constructs (for names, see Table 5.1) were transformed into *B. breve* UCC2003, *B. breve* JCM 7017 and *B. breve* NCFB 2258 in order to test if replication of these constructs was possible in bifidobacteria.

In order to determine if the observed non-replicative phenotype was due to restriction modification barriers, the constructs were passed through an *E. coli* methylating strain EC101-pNZ-M.Bbrll+Bbr111 (O'Connell Motherway *et al.*, 2009). Due to the fact that pBif contained a chloramphenicol marker as did the methylase strain EC101-pNZ-M.Bbrll+Bbr111 a new tet-resistant plasmid backbone, named pWSK29-tetW, was generated using the following primers Pwsk29Fsph1 and Pwsk29Rnco1, in the case of the plasmid backbone, and tetwAFL3F, tetwsph1R, tetwsph1F and tetwAFL3R incorporating *tetW* in both the forward and reverse orientation (constructs labelled S and A, respectively) (Supplementary Table S5.1). The individual DNA fragments encompassing $repA_{7017}$, $repA_{7017+1}$ and $repA_{7017+2}$ were removed by Sph1 restriction from pBif constructs, respectively, and ligated into SphI digested pWSK29-tetW. The ligation mixture was introduced into EC101 (Table 5.1) by electrotransformation, and transformants were then selected based on tetracycline and kanamycin resistance. The plasmid content of a number of Tet^r and

kn^r transformants was screened by restriction analysis and the integrity of positively identified clones was verified by sequencing. Following introduction of the resulting constructs into EC101-pNZ-M.Bbrll+Bbr111 to allow methylation of the plasmid DNA (Table 5.1) (O'Connell Motherway *et al.*, 2009), methylated plasmid DNA was obtained and used for transformation of *B. breve* UCC2003, *B. breve* JCM 7017 and *B. breve* NCFB 2258 in order to test if these constructs were capable of autonomous replication in Bifidobacteria.

Heterologous $RepA_{7017}$ (over)production

A 2 % inoculum of *L. lactis* NZ9000 harbouring plasmid pNZ-RepA-His in 25 ml of M17 broth supplemented with 0.5 % glucose was prepared, followed by incubation at 30^{o}C until an Optical Density (O.D. at wavelength 600 nm) of 0.5 was reached, at which point protein expression was induced by the addition of purified nisin (5 ng ml^{-1}) followed by continued incubation at 30^{o}C for 90 minutes. Cells were harvested, re-suspended in 10 mM Tris-HCl, and disrupted with glass beads in a mini-bead beater (BioSpec Products, Bartlesville, OK). Cellular debris was removed by centrifugation to produce a crude cell extract containing RepA-His.

Electrophoretic mobility shift assay (EMSA)

A DNA fragment, designated REP, representing $repA_{7017}$ and 702 bp of DNA immediately upstream of $repA_{7017}$ and including the promoter region, was prepared by PCR generating a 1752 bp product using an IRD800-labelled primer pair 124-repF and 124-repR (Supplementary table S5.2). EMSAs were performed essentially as described previously (Hamoen *et al.*, 1998). In all cases, the binding reactions were carried out in a final volume of 20 µl in the presence of poly[d(I-C)] in binding buffer (20 mM Tris-HCl [pH 7.0], 5 mM $MgCl_2$, 0.5 mM DTT, 1 mM EDTA, 200

mM KCl, 10 % glycerol). Varying amounts of crude cell extract of *L. lactis* NZ9000 harbouring plasmid pNZ-RepA-His (and expressing the $RepA_{7017}$ protein following nisin induction) and a fixed amount of DNA probe (0.1 pmol) were mixed on ice and subsequently incubated for 30 min at 37°C. Samples were loaded onto a 6 % non-denaturing Polyacrylamide (PAA) gel prepared in TAE buffer (40 mM Tris acetate [pH 8.0], 2 mM EDTA) and run in a 0.5-to-2.0 x gradient of TAE at 100 V for 90 minutes in an Atto Mini PAGE system (Atto Bioscience and Biotechnology, Tokyo, Japan). Signals were detected using Odyssey Infrared Imaging System (Li-Cor Biosciences UK Ltd, Cambridge, UK) and captured using the supplied software (Odyssey V3.0).

Primer extension analysis

Total RNA was isolated from *B. breve* UCC2003 pAM5-repAp, grown in mMRS supplemented with 1 % ribose to early exponential phase, using the Macaloid method (Kuipers *et al.*, 1993). RNA samples were treated with RNase-free DNase (Ambion). Primer extension was performed by annealing 1 pmol of IRD800 synthetic 18-mer oligonucleotides to 20 µg of RNA as previously described (Ventura *et al.*, 2005) using primers 124-REP-PE-R1 and 124-REP-PE-R2 (Supplementary Table S5.2). Sequence ladders of the presumed $repA_{7017}$ promoter region, which were run alongside the primer extension products, were produced using primers PE-DNA-F-124 and PE-DNA-R-124 (Supplementary Table S5.1) and employing the DNA cycle Sequencing Kit (Jena Bioscience, Loebstedter Strasse 80, Germany). Separation was achieved on a 6.5 % Li-Cor Matrix KB Plus acrylamide gel. Signal detection and image capture was performed by means of a Li-cor sequencing instrument (Li-Cor Biosciences).

5.4 RESULTS AND DISCUSSION

Identification of the bifidobacterial megaplasmid

We recently conducted a comparative genome analysis of representative members of the *B. breve* species, which revealed that *B. breve* JCM 7017 harbours a megaplasmid, designated pBb7017, with a size of over 190 Kb (Bottacini *et al.*, unpublished results). The presence and size of this plasmid in *B. breve* JCM 7017 was confirmed by pulsed field gel electrophoresis (Fig. 5.2). Sequence analysis of this plasmid identified two genes, *pl7017_0124* and *pl7017_0146*, designated here as $repA_{7017}$ and $repB_{7017}$, respectively, whose encoded products represent putative replication proteins (see below), suggesting that the pBb7017 plasmid is a co-integrate of two smaller plasmids. Further evidence for this apparent co-integration event can be derived from the evidently unevenly distributed GC content across the plasmid sequence. The plasmid is composed of two segments of differing GC content, where one segment, encompassing co-ordinates 119,440 through to 230,439, possesses an average GC-content of 60.1 %, while the other segment, encompassing the remainder of the plasmid, possesses an average GC-content of 55.3 % (Fig.5.3). Since this plasmid encodes functions that sustain stable replication of a very large piece of DNA, we wanted to identify and characterise these replication functions in order to create an *E. coli*-bifidobacterial shuttle and cloning vector for large DNA fragments.

In silico analysis of the $repA_{7017}$ and $repB_{7017}$ DNA regions

The sequences of the predicted replication proteins encoded by $repA_{7017}$ and $repB_{7017}$ were used to construct a neighbour joining phylogenetic tree (Saitou & Nei, 1987) displaying sequence similarities between $RepA_{7017}$ and $RepB_{7017}$, and a number of selected and mostly putative replication proteins (Fig. 5.4). Due to the paucity of

information regarding the majority of plasmid replication proteins we were not able to show direct and significant similarity to a characterised replication protein known to support either theta-type or rolling circle-type replication. However, the results of the phylogenetic tree showed that $RepA_{7017}$ and $RepB_{7017}$ both exhibit similarity to replication proteins that are predicted to replicate via the theta-type mode. Such replication proteins include those encoded by pCRY4, a 48 Kb plasmid isolated from *Corynebacterium glutamicum* (Tauch *et al*., 2002; Tauch *et al*., 2003), the 29 Kb plasmid pET44827 harboured by *Corynebacterium aurimucosum* ATCC 700975 (Trost *et al.*, 2010) and pCIBA089, a 2.1 Kb plasmid isolated from *Bifidobacterium asteroides* (Cronin *et al*., 2007) (Fig.5.4).

It is generally accepted that rolling circle-type plasmids are small in size, ranging approximately from 1.3 Kb to 10 Kb (Khan, 1997). However, plasmids that replicate by means of theta-type replication functions can represent both large and small plasmids, for example the 3.8 kb plasmid pWV02 from *L. lactis* replicates via the theta-type mechanism (Kiewiet *et al.*, 1993), and the 13 kb plasmid pRV500, isolated from *Lactobacillus sakei* RV332 (Alpert *et al.*, 2003). The putative replication protein $RepA_{7017}$ displays significant similarity to deduced replication proteins of a series of very large, uncharacterised plasmids, for example the 86.2 Kb pCHA1 plasmid from *Corynebacterium halotolerans* (Ruckert *et al.*, 2012), plasmid pRMAR01 (NC_013502) from *Rhodothermus marinus* DSM 4252 with a size of 125.133 Kb, and the 14.3 Kb plasmid pKW4 from *Corynebacterium jeikeium* K411 (Tauch *et al*., 2005).

Analysis of the DNA region preceding $repA_{7017}$ (schematically depicted in Fig. 5.5) allowed the detection of a high A+T rich region of 72.2 %, consisting of three imperfect direct repeats (Fig. 5.6A), a set of putative iterons (see below) and an

imperfect inverted repeat, features that have frequently been identified within the origin region of other theta-replicating plasmids (del Solar *et al.*, 1998). The putative iterons are represented by 8 nearly perfect, direct repeat sequences (Fig. 5.6B), which may be required for plasmid replication as such sequences are known to play a crucial role in initiation of replication. No sequences were detected which match the consensus sequence of the DnaA box (5'-TTATCCACA-3') (Weigel *et al.*, 1997), and which are commonly found within the origin of replication of theta-replicating plasmids (del Solar *et al.*, 1998). The presumed origin of replication structure upstream of $repA_{7017}$ resembles that, among others, of the putative theta-replicating bifidobacterial plasmid pCIBA089 (Cronin *et al.*, 2007) and that of the pSymb megaplasmid of *Rhizobium meliloti* (Margolin & Long, 1993). Upon investigation of the region preceding the replication gene $repB_{7017}$ we noticed that DNA sequences commonly associated with such replicons, such as direct and indirect repeats, were absent (data not shown). Therefore it is possible that $repB_{7017}$ is non-functional as a replication protein, and it may thus be that replication of pBb7017 is solely dependent on the replication functionalities provided by $repA_{7017}$ and associated sequences. For this reason we chose to further characterise $repA_{7017}$.

RepA$_{7017}$ binds to the repA$_{7017}$ promoter region

Bifidobacterial and other plasmid-encoded replication proteins have been shown to bind to repeat sequences upstream of their corresponding coding region; for example, Rep89 binds to direct repeats in the putative origin of replication of the putative theta-replicating plasmid pCIAB089 from *B. asteroides* (Cronin *et al.*, 2007). Examples in other bacteria would be that of predicted theta-replicating plasmids pPS10 from *Pseudomonas savastanoi* (de Viedma *et al.*, 1995), R2K from *E. coli* (Toukdarian *et al.*, 1996) and pSK41 from *Staphylococcus aureus* (Liu *et al.*,

2012). In order to establish if RepA$_{7017}$ can interact with the identified repeat sequences within the presumed promoter region upstream of the *repA$_{7017}$* gene we first cloned the *repA$_{7017}$* gene in the nisin-inducible vector pNZ8048 resulting in plasmid pNZ-repA$_{7017}$-His (see Materials and Methods). Following nisin induction of *L. lactis* NZ9000 pNZ-repA$_{7017}$-His, the recombinant protein was purified from the cell extract using metal affinity chromatography. However, purified RepA$_{7017}$ protein did not exhibit any binding activity, which is not uncommon when (replication) proteins of bifidobacterial origin are expressed in and purified from a heterologous host (Cronin *et al.*, 2007; O'Connell Motherway *et al.*, 2011; Pokusaeva *et al.*, 2011).

Therefore, instead of using the purified RepA$_{7017}$ protein, a crude cell extract of *L. lactis* NZ9000 pNZ-repA$_{7017}$-His was used for EMSA experiments. This approach allowed us to clearly demonstrate that RepA$_{7017}$ binds to the IRD800-labelled DNA fragment REP which encompasses a 1752 bp promoter-containing region (see below) upstream of *repA$_{7017}$* and the *repA$_{7017}$* gene (Fig. 5.7 (i)). Since crude cell extract of pNZ8048 (empty vector not expressing RepA$_{7017}$) did not exhibit binding towards the REP fragment mentioned above we conclude that RepA$_{7017}$ binds to sequences within the *repA$_{7017}$* promoter region (Fig. 5.7(ii)). DNA fragments that contained eight direct repeats present in the promoter region split into two sections were not affected in their mobility by RepA$_{7017}$ indicating that this protein seems to require all four for binding (results not shown). This result is in agreement with a previous report whereby the presumed replication protein Rep89 associated with the *B. asteroides* plasmid pCIBA089 was unable to bind to individual direct repeat sequences (Cronin *et al.*, 2007).

Identification of the repA$_{7017}$ transcription start site

In order to determine the transcription start site of the presumed *repA$_{7017}$* promoter, primer extension analysis was performed using RNA extracted from pAM5-repAp (pAM5 containing the presumed promoter region preceding *repA$_{7017}$*) from *B. breve* UCC2003 grown in Modified Rogosa medium containing 1 % ribose as the sole carbohydrate source. A single extension product was identified for the promoter region, corresponding to a position which is 71 nucleotides 5′ to the predicted translational start site for the *repA$_{7017}$* gene (Fig. 5.8). Analysis of the sequences upstream of this transcriptional start point within the promoter region of *repA$_{7017}$* revealed potential promoter recognition sequences resembling consensus −10 and −35 hexamers (Fig. 5.8). Interestingly, primer extension attempts using the wild type *B. breve* JCM 7017 strain which naturally harbours the mega plasmid proved unsuccessful. This may be due to low transcriptional levels of *repA$_{7017}$* as a result of negative transcriptional autoregulation by RepA$_{7017}$, a common phenomenon, which for example is found in pPS10 (de Viedma *et al.*, 1995). The experimental set-up we took for the primer extension experiment uncoupled such negative autoregulation of the *repA$_{7017}$* promoter (as this promoter was cloned separate from the *repA$_{7017}$* gene).

Attempts to construct shuttle vectors pbif-repA$_{7017}$, pbif-repA$_{7017}$+1 and pbif-repA$_{7017}$+2

Our presumption was that *repA$_{7017}$* and surrounding sequences are required for replication of the *B. breve* JCM 7017 megaplasmid pBb7017. In order to determine if *repA$_{7017}$* under the control of its own promoter and presumed origin of replication, could support replication in bifidobacteria a series of constructs were made using the pBif vector (Table 5.1). This vector was chosen as it replicates via a pSC101 *E. coli*

low copy number replicon and allows the cloning of large fragments in *E. coli* (Cohen & Chang, 1977). Our objective was to use this vector to clone the presumptive replication functions of the pBb7017 megaplasmid in order to create an *E. coli*/*Bifidobacterium* shuttle vector that would be suitable for cloning of large DNA fragments in order to express complex functions, for example those required for surface polysaccharide production (Fanning *et al.*, 2012; Tao & Zhang, 1998). Plasmids pbif-repA7017, pbif-repA7017+1 and pbif-repA7017+2 were first constructed in *E. coli* (see Materials and Methods). These constructs were then methylated using the EC101-pNZ-M.Bbrll+Bbr111 methylation strain (Table 5.1) (O'Connell Motherway *et al.*, 2009) in order to overcome possible restriction modification systems, and attempts were then made to introduce these plasmids into *B. breve* UCC2003, *B. breve* JCM 7017 or *B. breve* NCFB 2258 by electroporation. Unfortunately, no transformants were obtained for any of these strain-plasmid combinations. Electroporation with the pBC1.2-control plasmid (Table 5.1) demonstrated that the prepared cells for each strain were indeed competent. These results demonstrate that apparently not all replication functions were present on the three pBif derivatives or that these replication functions were dysfunctional.

Further cloning attempts will have to be made in order to determine if additional sequences are required for replication using the pBb7017 replication function (or at least those associated with $repA_{7017}$). Previously, bifidobacterial replication functions have been cloned and have been used successfully to construct *E.coli*/bifidobacterial shuttle vectors (Klijn *et al.*, 2006; Missich *et al.*, 1994; Rossi *et al.*, 1996; Rossi *et al.*, 1998; Sangrador-Vegas *et al.*, 2007; Shkoporov *et al.*, 2008). The use of a megaplasmid has been described for the creation of an *E. coli /Rhizobium meliloti* shuttle vector mini-pSym-b (Margolin & Long, 1993). In this particular study a 4 Kb

DNA fragment active as an autonomously replicating sequence from the *Rhizobium meliloti* megaplasmid designated pSym-b was isolated by selecting for sequences that allowed a normally non-replicative pBR322 derivative to replicate in *R. meliloti*. The 4kb fragment contained several sequence motifs common to other replication origins, such as an AT-rich region, three potential DnaA binding sites, a potential 13-mer sequence, and several groups of short direct repeats. The resulting *E. coli/R. meliloti* shuttle vector mini-pSym-b also replicated in the closely related strain *Agrobacterium tumefaciens*, but only in strains carrying the pSym-b megaplasmid, suggesting that a megaplasmid-encoded trans-acting factor is required for replication.

5.5 CONCLUDING REMARKS

This is the first study regarding the dissection of a megaplamid from a *Bifidobacterium* strain of human origin. Although the precise replication functions associated with megaplasmid pBb7017 from *B. breve* JCM 7017 have yet to be deciphered, this study provides significant data as regards the regulation of the $repA_{7017}$ promoter, and interactions between the $RepA_{7017}$ and its promoter region. Based on the *in silico* analysis presented in this study it is likely that pBb7017 replicates via a theta-type mechanism. One may further speculate that pBb7017 replicates in a manner that is similar to that of the *E. coli* chromosome (Bramhill & Kornberg, 1988). In the latter case the replication or Rep protein, binds to a series of direct repeats also known as iterons to form an initiation complex at the origin. In the case of *E. coli* chromosomal replication, this complex, in association with DnaA, allows the transfer of DnaB-DnaC to the replication origin, resulting in the opening of an AT-rich region, thereby initiating leading strand synthesis. The iteron-containing region (or even the complete plasmid sequence) of pBb7017 does not appear to contain DnaA boxes, and it is therefore possible that DnaB-DnaC loading to the replication origin requires other factors, perhaps aided by some of the sequence features that were identified in close proximity of the $RepA_{7017}$-recognised iterons. Further studies will have to be performed in order to pinpoint the precise mode of replication for the megaplasmid pBb7017.

5.6 ACKNOWLEDGEMENTS

This chapter has emanated from research conducted with the financial support of Science Foundation Ireland (SFI) under Grant Numbers 07/CE/B1368, 08/SRC/B13404 and SFI/12/RC/2273. KJOC was supported by a postgraduate fellowship funded through the Tomar trust, while MOCM is a recipient of a HRB postdoctoral fellowship (Grant no. PDTM/20011/9). Francesca Bottacini for her bioinformatics input.

5.7 TABLES AND FIGURES

TABLE 5.1 Bacterial strains and plasmids used in this study

Strain/ plasmid	Relevant characteristics	Reference or source
E.coli strains		
EC101	Cloning host, repA+ kmr	(Law *et al.*, 1995)
EC101-pNZ-M.Bbrll+Bbr111	EC101 harbouring pNZ8048 derivative containing bbrllM and bbrlllM	(O'Connell Motherway *et al.*, 2009)
E. coli XL1-blue	(*supE44 hsdR17 recA1 gyrA96 thi relA1 lac* F′ [*proAB*$^{+}$ *lacI*^{q} *lacZΔM15* Tn*10*(Tet^{r})])	Stratagene
L. lactis strains		
NZ9000	MG1363 derivative, nisin-inducible overexpression host; pepN::nisRK	(de Ruyter *et al.*, 1996)
NZ9700	Nisin producing strain	(de Ruyter *et al.*, 1996)
B.breve strains		
UCC2003	Isolate from a nursling stool	UCC
NCFB 2258	Isolate from infant intestine	NCFB
JCM 7017	Isolate from human faeces	JCM
Plasmids		
pNZ-RepA-His	*repA*$_{7017}$ with his tag cloned downstream of nisin inducible promoter on pNZ8048	This study
pAM5-repAp$_{7017}$	repA$_{7017}$ promoter region cloned in PAM5 (pBC1-puC19-Tcr)	This study
pBif-repA$_{7017}$	repA$_{7017}$ cloned in pBif	This study
pBif-repA$_{7017}$ +1	repA$_{7017}$ +1cloned in pBif	This study
pBif-repA$_{7017}$ +2	repA$_{7017}$ +2 cloned in pBif	This study

Strain/ plasmid	Relevant characteristics	Reference or source
pwsk29-tetwAFL	Tet^r, αlacZ,low copy number	This study
pwsk29-tetwSPH	Tet^r, αlacZ,low copy number	This study
pwsk29-tetw(a)-repA$_{7017}$	124 cloned in pwsk29-tetwAFL	This study
pwsk29-tetw(a)- repA$_{7017}$ +1	123+124 cloned in pwsk29-tetwAFL	This study
pwsk29-tetw(a)- repA$_{7017}$ +2	122-124 cloned in pwsk29-tetwAFL	This study
pwsk29-tetw(s)- repA$_{7017}$	124 cloned in pwsk29-tetwSPH	This study
pwsk29-tetw(s)- repA$_{7017}$ +1	123+124 cloned in pwsk29-tetwSPH	This study
pwsk29-tetw(s)- repA$_{7017}$ +2	122-124 cloned in pwsk29-tetwSPH	This study
pAM5	pBC1-puC19-Tc^r	(Alvarez-Martin *et al.*, 2007)
pNZ8048	Cm^r; nisin-inducible translational fusion vector	(de Ruyter *et al.*, 1996)
pBif	Ap^r, Cm^r , 6.5 kbp	APC-University College Cork
pBc1.2	pBC1 cloned in in pBif Cm^r	(Alvarez-Martin *et al.*, 2007)
pWSK29	Ap^r α*lac*Z,b low copy number	(Wang & Kushner, 1991)

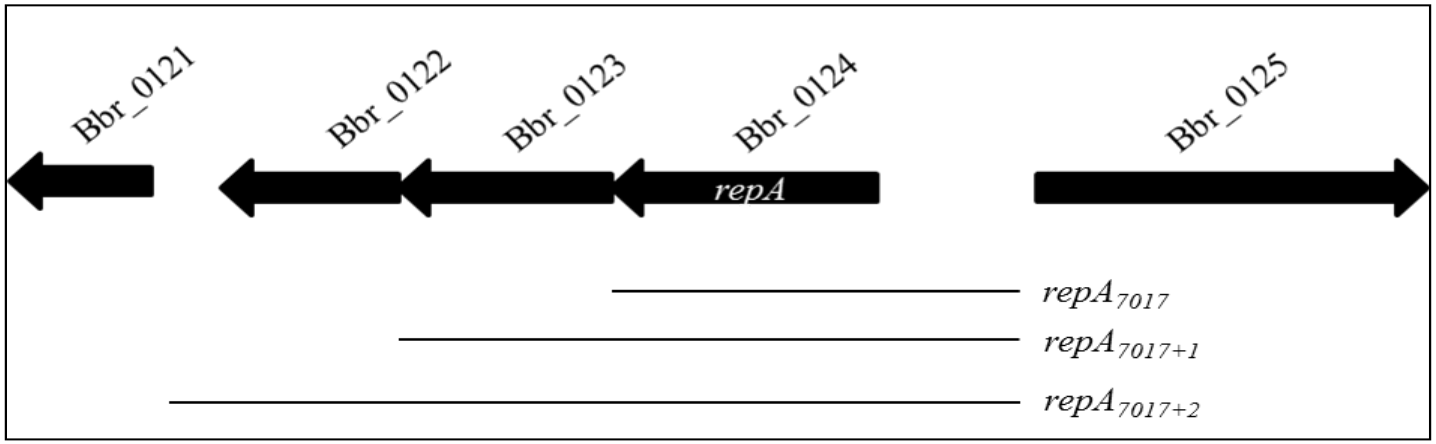

Figure 5.1 Schematic representation of cloned DNA fragments containing *repA*$_{7017}$, *repA*$_{7017+1}$, and *repA*$_{7017+2}$ including upstream and, in the case of *repA*$_{7017+2}$, downstream regions. Cloned fragments indicated by corresponding black line.

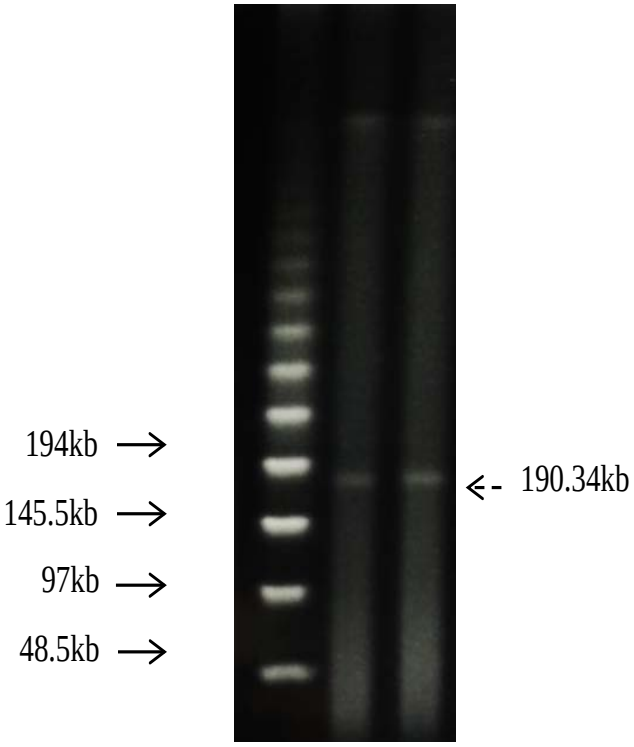

Figure 5.2 Pulse field gel electrophoresis of pBb7017, size represented at 190.34kb (both lanes). Dash type arrow indicates the S1 nuclease linearised megaplasmid band. Black compound type arrows indicate DNA size standards. Slight smearing indicates degradation due to S1 nuclease.

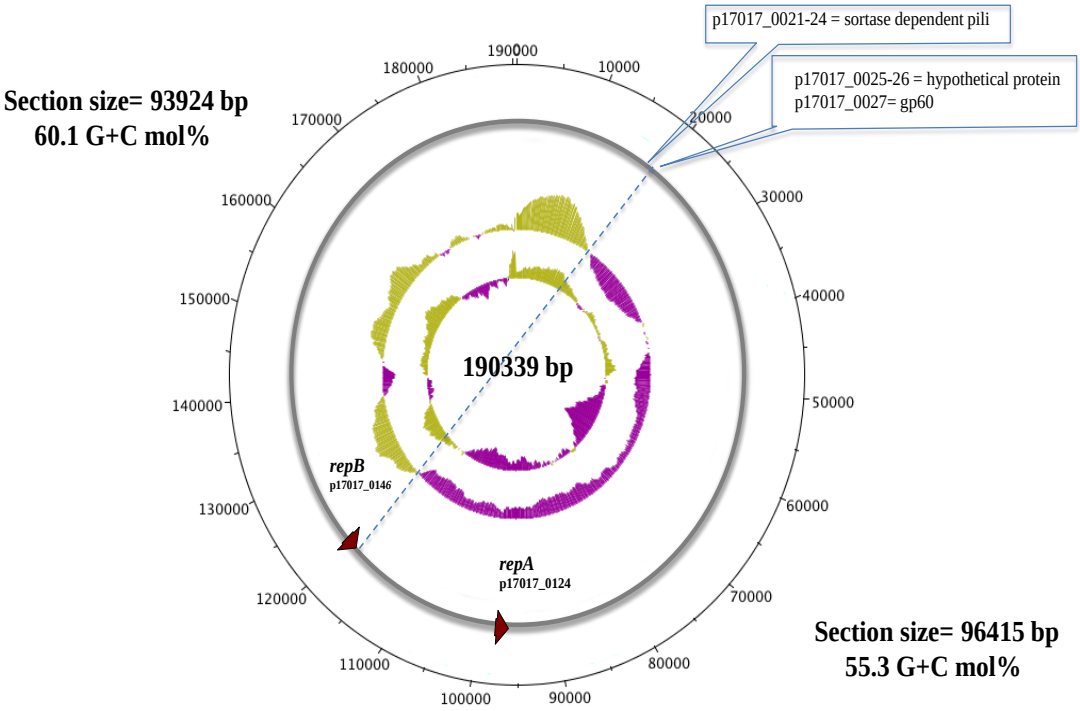

Figure 5.3 Schematic representation of pBb7017 plasmid, the inner to the outer circles represent GC skew, GC content, Replication genes (in red) and base position on the outer most circle. Dashed blue line represents putative dissection points of plasmid co-integration.

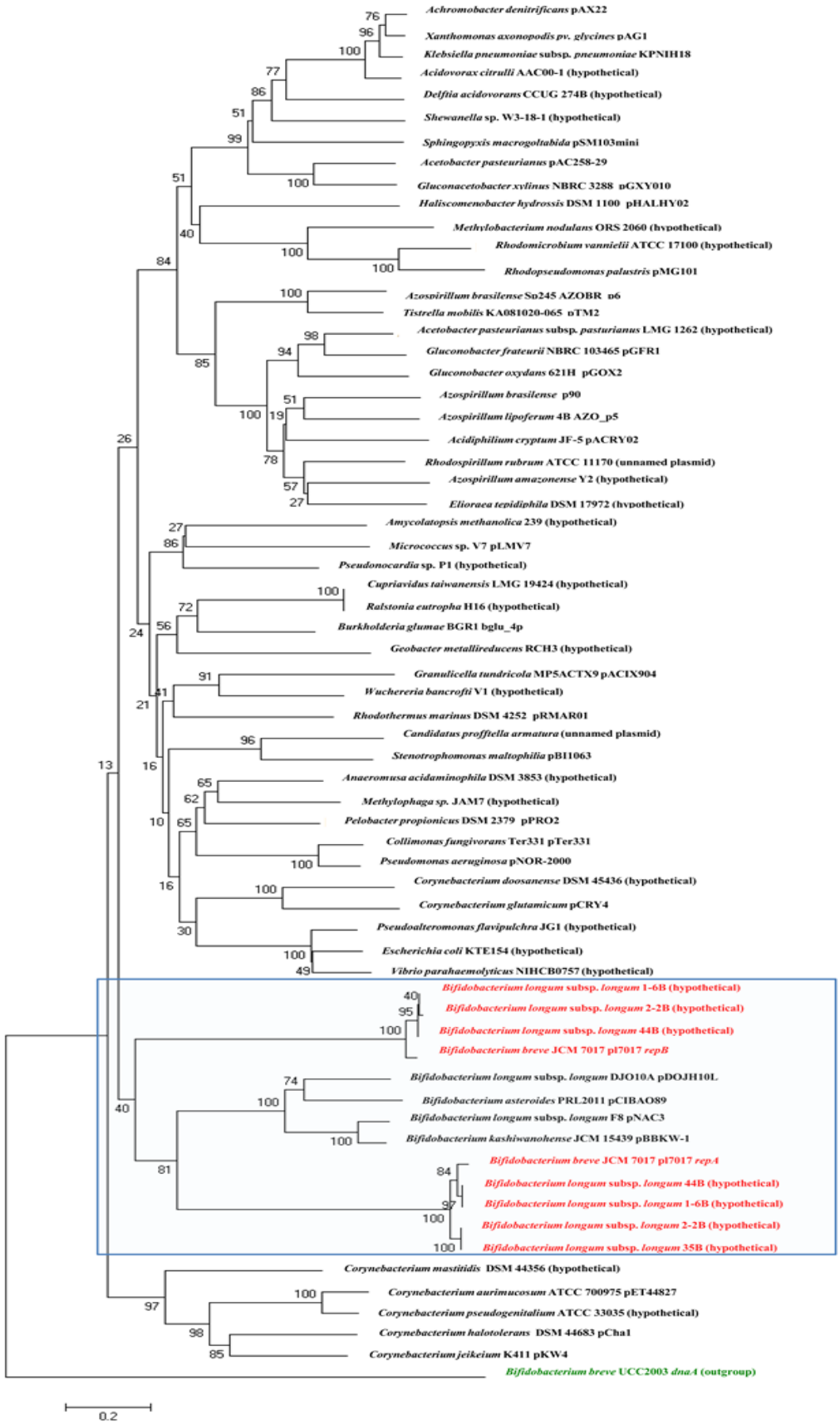
Achromobacter denitrificans pAX22
Xanthomonas axonopodis pv. glycines pAG1
Klebsiella pneumoniae subsp. pneumoniae KPNIH18
Acidovorax citrulli AAC00-1 (hypothetical)
Delftia acidovorans CCUG 274B (hypothetical)
Shewanella sp. W3-18-1 (hypothetical)
Sphingopyxis macrogoltabida pSM103mini
Acetobacter pasteurianus pAC258-29
Gluconacetobacter xylinus NBRC 3288 pGXY010
Haliscomenobacter hydrossis DSM 1100 pHALHY02
Methylobacterium nodulans ORS 2060 (hypothetical)
Rhodomicrobium vannielii ATCC 17100 (hypothetical)
Rhodopseudomonas palustris pMG101
Azospirillum brasilense Sp245 AZOBR p6
Tistrella mobilis KA081020-065 pTM2
Acetobacter pasteurianus subsp. pasturianus LMG 1262 (hypothetical)
Gluconobacter frateurii NBRC 103465 pGFR1
Gluconobacter oxydans 621H pGOX2
Azospirillum brasilense p90
Azospirillum lipoferum 4B AZO_p5
Acidiphilium cryptum JF-5 pACRY02
Rhodospirillum rubrum ATCC 11170 (unnamed plasmid)
Azospirillum amazonense Y2 (hypothetical)
Elioraea tepidiphila DSM 17972 (hypothetical)
Amycolatopsis methanolica 239 (hypothetical)
Micrococcus sp. V7 pLMV7
Pseudonocardia sp. P1 (hypothetical)
Cupriavidus taiwanensis LMG 19424 (hypothetical)
Ralstonia eutropha H16 (hypothetical)
Burkholderia glumae BGR1 bglu_4p
Geobacter metallireducens RCH3 (hypothetical)
Granulicella tundricola MP5ACTX9 pACIX904
Wuchereria bancrofti V1 (hypothetical)
Rhodothermus marinus DSM 4252 pRMAR01
Candidatus protofella armatura (unnamed plasmid)
Stenotrophomonas maltophilia pBI1063
Anaeromusa acidaminophila DSM 3853 (hypothetical)
Methylophaga sp. JAM7 (hypothetical)
Pelobacter propionicus DSM 2379 pPRO2
Collimonas fungivorans Ter331 pTer331
Pseudomonas aeruginosa pNOR-2000
Corynebacterium doosanense DSM 45436 (hypothetical)
Corynebacterium glutamicum pCRY4
Pseudoalteromonas flavipulchra JG1 (hypothetical)
Escherichia coli KTE154 (hypothetical)
Vibrio parahaemolyticus NIHCB0757 (hypothetical)
Bifidobacterium longum subsp. longum 1-6B (hypothetical)
Bifidobacterium longum subsp. longum 2-2B (hypothetical)
Bifidobacterium longum subsp. longum 44B (hypothetical)
Bifidobacterium breve JCM 7017 pI7017 repB
Bifidobacterium longum subsp. longum DJO10A pDOJH10L
Bifidobacterium asteroides PRL2011 pCIBAO89
Bifidobacterium longum subsp. longum F8 pNAC3
Bifidobacterium kashiwanohense JCM 15439 pBBKW-1
Bifidobacterium breve JCM 7017 pI7017 repA
Bifidobacterium longum subsp. longum 44B (hypothetical)
Bifidobacterium longum subsp. longum 1-6B (hypothetical)
Bifidobacterium longum subsp. longum 2-2B (hypothetical)
Bifidobacterium longum subsp. longum 35B (hypothetical)
Corynebacterium mastitidis DSM 44356 (hypothetical)
Corynebacterium aurimucosum ATCC 700975 pET44827
Corynebacterium pseudogenitalium ATCC 33035 (hypothetical)
Corynebacterium halotolerans DSM 44683 pCha1
Corynebacterium jeikeium K411 pKW4
Bifidobacterium breve UCC2003 dnaA (outgroup)
0.2

Figure 5.4 Neighbour joining phylogenetic tree of replication genes *repA*$_{7017}$ and *repB*$_{7017}$ from pBb7017 as compared to other bacteria. Red indicates most related, these are indicated by a blue box. Green text represents the outlier.

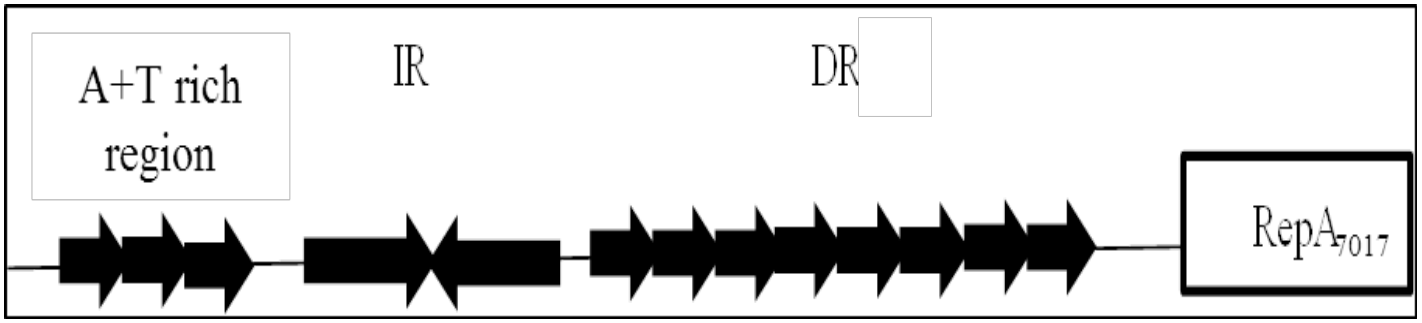

Figure 5.5 Structural features of pBb7017 representing the high A+T rich region, and inverted repeat (IR) and eight direct repeats (DR) followed by the *repA*$_{7017}$ protein gene. See also Figure 5.8 for further sequence details.

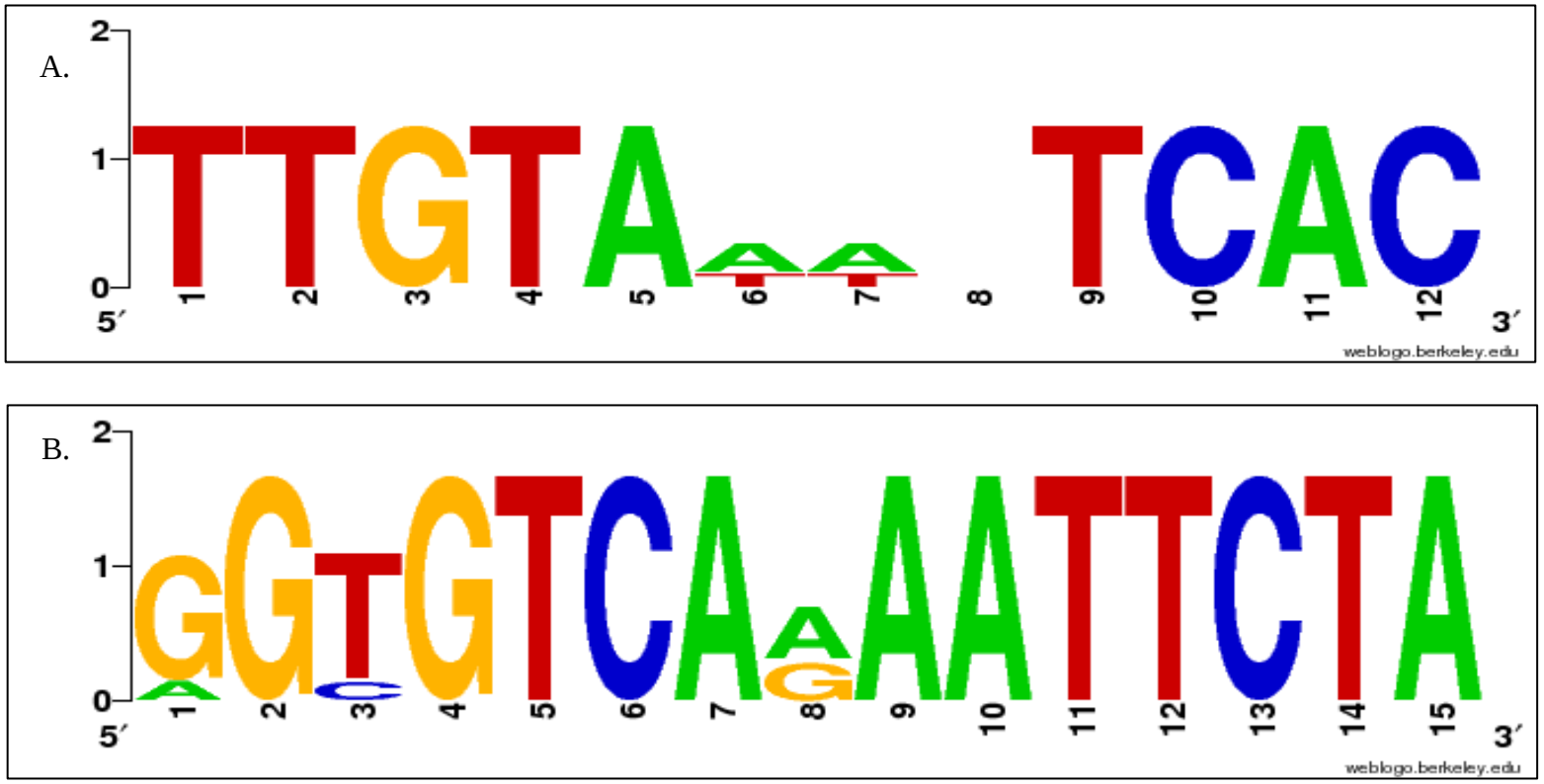

Figure 5.6 Web logo representation of predicted consensus repeat sequences; Panel A: Comparison of three direct repeats in the high A+T rich region; Panel B: Consensus derived from the eight direct repeat sequences present in the *repA*$_{7017}$ promoter sequence.

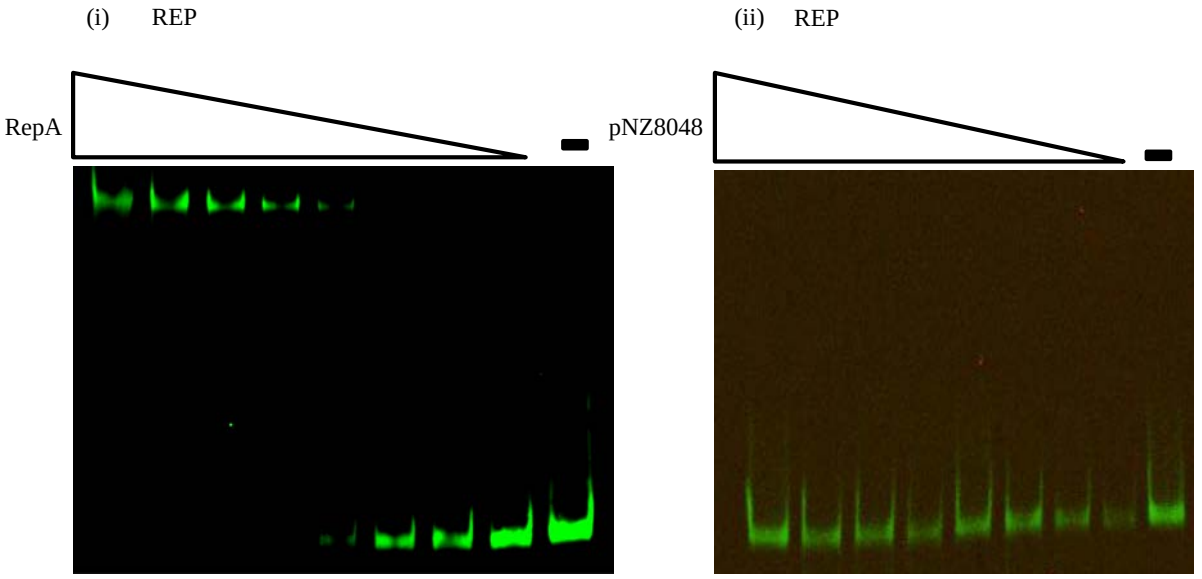

Figure 5.7 EMSAs showing interaction of crude cell extract of nisin induced NZ9000 (i) pNZ-RepA-His or (ii) pNZ8048 with DNA fragment encompassing promoter region preceding *repA*$_{7017}$.

-360 -340

GAACGGATTGCATCCCTGTAAGTCGCAGGTCGAAGTATGAGGGAGATT

-320 -300

AAGAGAGAGATTGTAAAGTCACTTGTAATATCACTTGTATATTCACTT

-280 -260 -240

GGGGGTGCCTCTTAGCCCTACTGCCATAAGGGCTGAGAGGCCATGTCG

-220 -200

GTGTCAAAATTCTAAGAGTGGGTGTCAGAATTCTAAGCGAGGGTGTCA

-180 -160 -140

AAATTCTATAAACCGGTGTCAAAATTCTAAGAAAAAAACGCTCCGTTT

-120 -100

TGCGACTTAATGGGGTGGAATATTCGATTCCGTCAAAAGGGTGTCAGA

-80 -60

ATTCTAAGGAAAGCGTCAAAATTCTAAGATAGGTGTCAATATTCTAAA

-40 **-35** **-10** +1

TTTGGTG**TCAGAA**TTCTAAGAAAACGAGG**TAGGAT**AGCAGAGTCGAAA

+20 +40

AAAAGAAGATACCCGCACCGTTTGCGCCGGCCGGGTATCTGGCAATCG

+60

CT**GAAAGG**TAGCGACCGCATG

RBS

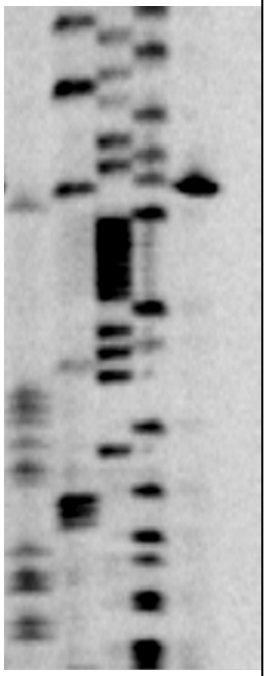

Figure 5.8 Boldface type and underlining indicate the −10 and −35 hexamers as deduced from the primer extension results and ribosomal binding site (RBS); the transcriptional start sites (TSS) are indicated by asterisks; Blue arrows three imperfect direct repeats which represent high A+T rich region; Red arrows indicate eight near perfect direct repeats associated with $RepA_{7017}$ binding; Black compound arrows represent an imperfect inverted repeat.

5.8 SUPPLEMENTARY MATERIAL

Table S5.1 Oligonucleotide primers used in this study

Name	Sequence	Product (bp)
124CTERMFNCO1	TGCACG**CCATGG**CCACCAATAATAACGCCATCGC	
124 CTERMRXBA1	CTATGC**TCTAGA**TCAGTGATGGTGATGGTGATGGTGATGGTGATGACCCAACAAAACCCGAGAC	1047
repApXba1F	CTATGC**TCTAGA**TGAGAGGCCATGTCGGTG	
repApSph1R	CAGTCA**GCATGC**CGAATCATGGACTGCACGT	539
Pwsk29Fsph1	TAACAA**GCATGC**GCATGGTCTATCAGGGCG	
Pwsk29Rnco1	AAAGGA**CCATGG**CTCATATATATACTTTAGATTG	2624
tetwAFL3F	TCAGCT**ACATGT**ATGCTCATGTACGGTAAG	
tetwsph1R	GCGACG**GCATGC**CCGACTGTTACTTTTTACAGTCGGTTTTCTAATGTCACCATTACCTTCTGAAACATA	2574
tetwsph1F	TCAGCT**GCATGC**ATGCTCATGTACGGTAAG	
tetwAFL3R	GCGACG**ACATGT**CCGACTGTACTTTTTACAGTCGGTTTTCTAATGTCACCATTACCTTCTGAAACATA	2574
124FSPH1-REP	CCGTCA**GCATGC**CTGCTAGTTCTTCTTCGTGCCATAC	
124RSPH1-REP	CAGTCA**GCATGC**CGAGGGTTTTCATGATGGC	1829
123+124FSPH1-REP	CCGTCA**GCATGC**CTGCTAGTTCTTCTTCGTGCCATAC	
123+124RSPH1-REP	CCGTCA**GCATGC**GTTGACCGTGATGTACTGGCTC	2730
122-124FSPH1-REP	CCGTCA**GCATGC**CTGCTAGTTCTTCTTCGTGCCATAC	
122-124RSPH1-REP	CTGTCA**GCATGC**GACTTTGAAGTTGATGATG	3560

[a]Restriction sequence is indicated in bold

Table S5.2 IRD800 labelled primers used in this study

Name	IRD800 labelled Primer	Product (bp)
124-rep F	CTGCTAGTTCTTCTTCGTGCCATAC	
124-rep R	TCAACCCAACAAAACCCGAGAC	1752
PE-DNA-F-124	TGAGAGGCCATGTCGGTG	
PE-DNA-R-124	CATTCCTCGGGCCGGCT	485
124-REP-PE-R1	GCGATGGCGTTATTATTGGT	
124-REP-PE-R2	CGAGGTACTGGGTTTTGACA	

5.9 REFERENCES

Almeida, C. C., Lorena, S. L. S., Pavan, C. R., Akasaka, H. M. I. & Mesquita, M. A. (2012). Beneficial Effects of Long-Term Consumption of a Probiotic Combination of *Lactobacillus casei Shirota* and *Bifidobacterium breve* Yakult May Persist After Suspension of Therapy in Lactose-Intolerant Patients. *Nutrition in Clinical Practice* **27**, 247-251.

Alpert, C. A., Crutz-Le Coq, A. M., Malleret, C. & Zagorec, M. (2003). Characterisation of a theta-type plasmid from *Lactobacillus sakei*: a potential basis for low-copy-number vectors in *lactobacilli*. *Applied and Environmental Microbiology* **69**, 5574-5584.

Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. (1990). Basic local alignment search tool. *Journal of Molecular Biology* **215**, 403-410.

Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J., Zhang, Z., Miller, W. & Lipman, D. J. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Research* **25**, 3389-3402.

Alvarez-Martin, P., O'Connell-Motherway, M., van Sinderen, D. & Mayo, B. (2007). Functional analysis of the pBC1 replicon from *Bifidobacterium catenulatum* L48. *Applied Microbiology and Biotechnology* **76**, 1395-1402.

Bai, Y. L., Yang, Z. L., Qiao, M. Q., Zhang, X. M., Zhou, J. & Gao, C. C. (2003). The action of S1 nuclease and a cloning strategy for microcircular DNAs. *Chinese Journal of Biotechnology* **19**, 240-243.

Beard, P., Morrow, J. F. & Berg, P. (1973). Cleavage of circular, superhelical simian virus 40 DNA to a linear duplex by S1 nuclease. *Journal of Virology* **12**, 1303-1313.

Bordoni, A., Amaretti, A., Leonardi, A., Boschetti, E., Danesi, F., Matteuzzi, D., Roncaglia, L., Raimondi, S. & Rossi, M. (2013). Cholesterol-lowering probiotics: *in vitro* selection and *in vivo* testing of bifidobacteria. *Applied Microbiology and Biotechnology* **97**, 8273-8281.

Bramhill, D. & Kornberg, A. (1988). Duplex opening by dnaA protein at novel sequences in initiation of replication at the origin of the *E. coli* chromosome. *Cell* **52**, 743-755.

Claesson, M. J., Li, Y., Leahy, S., Canchaya, C., van Pijkeren, J.P., Cerdeño-Tárraga, A.M., Parkhill, J., Flynn, S., O' Sullivan, G.C., Collins, J.K., Higgins, D., Shanahan, F., Fitzgerald, G.F., van Sinderen, D. & O'Toole PW (2006). Multireplicon genome architecture of *Lactobacillus salivarius*. *Proceedings of the National Academy of Sciences of the United States of America* **103**, 6718-6723.

Cohen, S. N. & Chang, A. C. Y. (1977). Revised Interpretation of the Origin of the pSC101 Plasmid. *Journal of Bacteriology* **132**, 734-737.

Corneau, N., Emond, E. & LaPointe, G. (2004). Molecular characterisation of three plasmids from *Bifidobacterium longum*. *Plasmid* **51**, 87-100.

Cronin, M., Knobel, M., O'Connell-Motherway, M., Fitzgerald, G. F. & van Sinderen, D. (2007). Molecular Dissection of a Bifidobacterial Replicon. *Applied and Environmental Microbiology* **73**, 7858-7866.

Cronin, M., Morrissey, D., Rajendran, S., El Mashad, S. M., van Sinderen, D., O'Sullivan, G. C. & Tangney, M. (2010). Orally administered bifidobacteria as vehicles for delivery of agents to systemic tumors. *Molecular therapy : the journal of the American Society of Gene Therapy* **18**, 1397-1407.

Cronin, M., Akin, A. R., Collins, S. A., Meganck. J., Kim, J.B., Baban, C.K., Joyce, S.A., van Dam, G.M., Zhang, N., van Sinderen, D., O' Sullivan, G.C., Kasahara, N., Gahan, C.G., Francis, K.P. & Tangney, M. (2012). High Resolution Bioluminescent Imaging for the Study of Bacterial Tumour Targeting. *PLoS ONE* **7**, e30940.

Dambekodi, P. C. & Gilliland, S. E. (1998). Incorporation of cholesterol into the cellular membrane of *Bifidobacterium longum*. *Journal of Dairy Science* **81**, 1818-1824.

de Ruyter, P. G., Kuipers, O. P., Beerthuyzen, M. M., van Alen-Boerrigter, I. & de Vos, W. M. (1996). Functional analysis of promoters in the nisin gene cluster of *Lactococcus lactis*. *Journal of Bacteriology* **178**, 3434-3439.

de Viedma, D. G., serrano-Loópez, A. & Diaz-Orejas, R. (1995). Specific binding of the replication protein of plasmid pPS10 to direct and inverted repeats is mediated by an HTH motif. *Nucleic Acids Research* **23**, 5048-5054.

del Solar, G., Giraldo, R., Ruiz-Echevarría, M. J., Espinosa, M. & Díaz-Orejas, R. (1998). Replication and control of circular bacterial plasmids. *Microbiology and Molecular Biology Reviews* **62**, 434-464.

Fanning, S., Hall, L. J., Cronin, M. Zomer, A., MacSharry, J., Goulding, D., O'Connell Motherway, M., Shanahan, F., Nally, K., Dougan, G. & van Sinderen, D. (2012). Bifidobacterial surface-exopolysaccharide facilitates commensal-host interaction through immune modulation and pathogen protection. *Proceedings of the National Academy of Sciences* **109**, 2108-2113.

Fukuda, S., Toh, H., Hase, K., Oshima, K., Nakanishi, Y., Yoshimura, K., Tobe, T., Clarke, J.M., Topping, D.L., Suzuki, T., Taylor, T.D., Itoh, K., Kikuchi, J., Morita, H., Hattori, M. & Ohno, H. (2011). Bifidobacteria can protect from enteropathogenic infection through production of acetate. *Nature* **469**, 543-547.

Germond, J. E., Vogt, V. M. & Hirt, B. (1974). Characterisation of the single-strand-specific nuclease S1 activity on double-stranded supercoiled polyoma DNA. *European Journal of Biochemistry* **43**, 591-600.

H.Tissier (1906). The treatment of intestinal infections by the method of transformation of the bacterial intestinal flora. *CR Soc Biol* **60**, 359-361.

Hamoen, L. W., Van Werkhoven, A. F., Bijlsma, J. J., Dubnau, D. & Venema, G. (1998). The competence transcription factor of *Bacillus subtilis* recognises short A/T-rich sequences arranged in a unique, flexible pattern along the DNA helix. *Genes & Development* **12**, 1539-1550.

Khan, S. A. (1997). Rolling-circle replication of bacterial plasmids. *Microbiology and Molecular Biology Reviews* **61**, 442-455.

Kiewiet, R., Bron, S., de Jonge, K., Venema, G. & Seegers, J. F. (1993). Theta replication of the lactococcal plasmid pWVO2. *Molecular Microbiology* **10**, 319-327.

Klijn, A., Moine, D., Delley, M., Mercenier, A., Arigoni, F. & Pridmore, R. D. (2006). Construction of a Reporter Vector for the Analysis of *Bifidobacterium longum* Promoters. *Applied and Environmental Microbiology* **72**, 7401-7405.

Kuipers, O. P., Beerthuyzen, M. M., Siezen, R. J. & de Vos, W. M. (1993). Characterisation of the nisin gene cluster nisABTCIPR of *Lactococcus lactis* . Requirement of expression of the nisA and nisI genes for development of immunity. *European Journal of Biochemistry* **216**, 281-291.

Law, J., Buist, G., Haandrikman, A., Kok, J., Venema, G. & Leenhouts, K. (1995). A system to generate chromosomal mutations in *Lactococcus lactis* which allows fast analysis of targeted genes. *Journal of Bacteriology* **177**, 7011-7018.

Lee, J. H. & O'Sullivan, D. J. (2006). Sequence analysis of two cryptic plasmids from *Bifidobacterium longum* DJO10A and construction of a shuttle cloning vector. *Applied and Environmental Microbiology* **72**, 527-535.

Li, Y., Canchaya, C., Fang, F., Raftis, E., Ryan, K. A., van Pijkeren, J.-P., van Sinderen, D. & O'Toole, P. W. (2007). Distribution of Megaplasmids in *Lactobacillus salivarius* and Other *Lactobacilli*. *Journal of Bacteriology* **189**, 6128-6139.

Liu, M. A., Kwong, S. M., Pon, C. K., Skurray, R. A. & Firth, N. (2012). Genetic requirements for replication initiation of the staphylococcal multiresistance plasmid pSK41. *Microbiology* **158**, 1456-1467.

Ma, Y., Luo, Y., Huang, X., Song, F. & Liu, G. (2012). Construction of *Bifidobacterium infantis* as a live oral vaccine that expresses antigens of the major fimbrial subunit (CfaB) and the B subunit of heat-labile enterotoxin (LTB) from enterotoxigenic *Escherichia coli*. *Microbiology* **158**, 498-504.

Margolin, W. & Long, S. R. (1993). Isolation and characterisation of a DNA replication origin from the 1,700-kilobase-pair symbiotic megaplasmid pSym-b of *Rhizobium meliloti*. *Journal of Bacteriology* **175**, 6553-6561.

Maze, A., O'Connell-Motherway, M., Fitzgerald, G. F., Deutscher, J. & van Sinderen, D. (2007). Identification and characterisation of a fructose phosphotransferase system in *Bifidobacterium breve* UCC2003. *Applied and Environmental Microbiology* **73**, 545-553.

Mierau, I., Leij, P., van Swam, I., Blommestein, B., Floris, E., Mond, J. & Smid, E. (2005). Industrial-scale production and purification of a heterologous protein in *Lactococcus lactis* using the nisin-controlled gene expression system NICE: The case of lysostaphin. *Microbial Cell Factories* **4**, 15.

Missich, R., Sgorbati, B. & LeBlanc, D. J. (1994). Transformation of *Bifidobacterium longum* with pRM2, a constructed *Escherichia coli-B. longum* shuttle vector. *Plasmid* **32**, 208-211.

Miyake, T., Watanabe, K., Watanabe, T. & Oyaizu, H. (1998). Phylogenetic analysis of the genus *Bifidobacterium* and related genera based on 16S rDNA sequences. *Microbiology and Immunology* **42**, 661-667.

Moon, G. S., Wegmann, U., Gunning, A. P., Gasson, M. J. & Narbad, A. (2009). Isolation and characterisation of a theta-type cryptic plasmid from *Bifidobacterium longum* FI10564. *Journal of Microbiology and Biotechnology* **19**, 403-408.

O'Connell Motherway, M., O'Driscoll, J., Fitzgerald, G. F. & Van Sinderen, D. (2009). Overcoming the restriction barrier to plasmid transformation and targeted mutagenesis in *Bifidobacterium breve* UCC2003. *Microbial Biotechnology* **2**, 321-332.

O'Connell Motherway, M., Fitzgerald, G. F. & van Sinderen, D. (2011). Metabolism of a plant derived galactose-containing polysaccharide by *Bifidobacterium breve* UCC2003. *Microbial Biotechnology* **4**, 403-416.

O'Riordan, K. & Fitzgerald, G. F. (1999). Molecular characterisation of a 5.75-kb cryptic plasmid from *Bifidobacterium breve* NCFB 2258 and determination of mode of replication. *FEMS Microbiology Letters* **174**, 285-294.

O' Shea, E.F., O' Connor, P.M., Raftis, E.J., O' Toole P.W., Stanton, C., Cotter, P.D., Ross, R.P. & Hill, C. Production of multiple bacteriocins from a single locus by gastrointestinal strains of *Lactobacillus salivarius*. *Journal of Bacteriology* **193,** 6973-6982.

Park, M. S., Lee, K. H. & Ji, G. E. (1997). Isolation and characterisation of two plasmids from *Bifidobacterium longum*. *Letters in Applied Microbiology* **25**, 5-7.

Pokusaeva, K., O'Connell-Motherway, M., Zomer, A., Macsharry, J., Fitzgerald, G. F. & van Sinderen, D. (2011). Cellodextrin utilisation by *Bifidobacterium breve* UCC2003. *Applied and Environmental Microbiology* **77**, 1681-1690.

Rossi, M., Brigidi, P., Gonzalez Vara y Rodriguez, A. & Matteuzzi, D. (1996). Characterisation of the plasmid pMB1 from *Bifidobacterium longum* and its use for shuttle vector construction. *Research in Microbiology* **147**, 133-143.

Rossi, M., Brigidi, P. & Matteuzzi, D. (1998). Improved cloning vectors for *Bifidobacterium* spp. *Letters in Applied Microbiology* **26**, 101-104.

Ruckert, C., Albersmeier, A., Al-Dilaimi, A., Niehaus, K., Szczepanowski, R. & Kalinowski, J. (2012). Genome sequence of the halotolerant bacterium *Corynebacterium halotolerans* type strain YIM 70093(T) (= DSM 44683(T)). *Standards in Genomic Sciences* **7**, 284-293.

Ruiz, L., Motherway, M. O., Lanigan, N. & van Sinderen, D. (2013). Transposon mutagenesis in *Bifidobacterium breve*: construction and characterisation of a Tn5 transposon mutant library for *Bifidobacterium breve* UCC2003. *PLoS One* **8**, e64699.

Rutherford, K., Parkhill, J., Crook, J., Horsnell, T., Rice, P., Rajandream, M. A. & Barrell, B. (2000). Artemis: sequence visualisation and annotation. *Bioinformatics* **16**, 944-945.

Saitou, N. & Nei, M. (1987). The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Molecular Biology and Evolution* **4**, 406-425.

Sambrook, J., Fritsch, E. F. & Maniatis, T. (1989). *Molecular cloning: a laboratory manual*: Cold Spring Harbor Laboratory.

Sangrador-Vegas, A., Stanton, C., van Sinderen, D., Fitzgerald, G. F. & Ross, R. P. (2007). Characterisation of plasmid pASV479 from *Bifidobacterium*

pseudolongum subsp. *globosum* and its use for expression vector construction. *Plasmid* **58**, 140-147.

Sgorbati, B., Scardovi, V. & Leblanc, D. J. (1982). Plasmids in the Genus *Bifidobacterium*. *Journal of General Microbiology* **128**, 2121-2131.

Shkoporov, A. N., Efimov, B. A., Khokhlova, E. V., Steele, J. L., Kafarskaia, L. I. & Smeianov, V. V. (2008). Characterisation of plasmids from human infant *Bifidobacterium* strains: sequence analysis and construction of E. coli-*Bifidobacterium* shuttle vectors. *Plasmid* **60**, 136-148.

Simpson, P. J., Stanton, C., Fitzgerald, G. F. & Ross, R. P. (2003). Genomic diversity and relatedness of bifidobacteria isolated from a porcine cecum. *Journal of Bacteriology* **185**, 2571-2581.

Sun, Z., Baur, A., Zhurina, D., Yuan, J. & Riedel, C. U. (2012). Accessing the Inaccessible: Molecular Tools for Bifidobacteria. *Applied and Environmental Microbiology* **78**, 5035-5042.

Takahata, M., Toh, H., Nakano, A., Takagi, M., Murakami, M., Ishii, Y., Takizawa, T., Tanabe, S. & Morita, H. (2013). Complete sequence analysis of two cryptic plasmids from *Bifidobacterium kashiwanohense* JCM 15439 (type strain) isolated from healthy infant feces. *Animal Science Journal* **85,** 158-63.

Tao, Q. & Zhang, H.-B. (1998). Cloning and stable maintenance of DNA fragments over 300 kb in *Escherichia coli* with conventional plasmid-based vectors. *Nucleic Acids Research* **26**, 4901-4909.

Tauch, A., Gotker, S., Puhler, A., Kalinowski, J. & Thierbach, G. (2002). The 27.8-kb R-plasmid pTET3 from *Corynebacterium glutamicum* encodes the aminoglycoside adenyltransferase gene cassette aadA9 and the regulated tetracycline efflux system Tet 33 flanked by active copies of the widespread insertion sequence IS6100. *Plasmid* **48**, 117-129.

Tauch, A., Puhler, A., Kalinowski, J. & Thierbach, G. (2003). Plasmids in *Corynebacterium glutamicum* and their molecular classification by comparative genomics. *Journal of Biotechnology* **104**, 27-40.

Tauch, A., Kaiser, O., Hain, T. Goesmann, A., Weisshaar, B., Albersmeier, A., Bekel, T., Bischoff, N., Brune, I., Chakraborty, T., Kalinowski, J., Meyer, F., Rupp, O., Schneiker, S., Viehoever, P. & Pühler, A. (2005). Complete Genome Sequence and Analysis of the Multiresistant Nosocomial Pathogen *Corynebacterium jeikeium* K411, a Lipid-Requiring Bacterium of the Human Skin Flora. *Journal of Bacteriology* **187**, 4671-4682.

Terzaghi, B. E. & Sandine, W. E. (1975). Improved medium for lactic *Streptococci* and their bacteriophages. *Applied Microbiology* **29**, 807-813.

Toukdarian, A. E., Helinski, D. R. & Perri, S. (1996). The Plasmid RK2 Initiation Protein Binds to the Origin of Replication as a Monomer. *Journal of Biological Chemistry* **271**, 7072-7078.

Trost, E., Gotker, S., Schneider, J. Schneiker-Bekel, S., Szczepanowski, R., Tilker, A., Viehoever, P., Arnold, W., Bekel, T., Blom, J., Gartemann, K.H., Linke, B., Goesmann, A., Pühler, A., Shukla, S.K. & Tauch, A. (2010). Complete genome sequence and lifestyle of black-pigmented *Corynebacterium aurimucosum* ATCC 700975 (formerly C. nigricans CN-1) isolated from a vaginal swab of a woman with spontaneous abortion. *BMC Genomics* **11**, 91.

Ventura, M., Zink, R., Fitzgerald, G. F. & van Sinderen, D. (2005). Gene Structure and Transcriptional Organisation of the dnaK Operon of *Bifidobacterium breve* UCC 2003 and Application of the Operon in Bifidobacterial Tracing. *Applied and Environmental Microbiology* **71**, 487-500.

Ventura, M., Canchaya, C., Tauch, A., Chandra, G., Fitzgerald, G. F., Chater, K. F. & van Sinderen, D. (2007). Genomics of *Actinobacteria*: Tracing the Evolutionary History of an Ancient Phylum. *Microbiology and Molecular Biology Reviews* **71**, 495-548.

Wang, R. F. & Kushner, S. R. (1991). Construction of versatile low-copy-number vectors for cloning, sequencing and gene expression in *Escherichia coli*. *Gene* **100**, 195-199.

Weigel, C., Schmidt, A., Ruckert, B., Lurz, R. & Messer, W. (1997). DnaA protein binding to individual DnaA boxes in the *Escherichia coli* replication origin, oriC. *The EMBO Journal* **16**, 6574-6583.

Wells, J. M., Wilson, P. W. & Le Page, R. W. (1993). Improved cloning vectors and transformation procedure for *Lactococcus lactis*. *The Journal of Applied Bacteriology* **74**, 629-636.

Chapter VI

General Discussion and Future Perspectives

Bifidobacterial genome sequence analysis has progressed significantly in recent years. There are currently 43 draft bifidobacterial genome sequences at different stages of completion in the database of National Centre of Biotechnology Information (NCBI; www.ncbi.nlm.nih.gov). This bifidobacterial genome sequence analysis has provided a wealth of primary information that has facilitated a better understanding as to how these bacteria survive in the gastro intestinal tract (GIT) and how they elicit purported beneficial activities towards their host (For a review see (Ventura *et al.*, 2014)). There has also been a steady rise in the number of scientific papers that reported on the sequencing and analysis of bifidobacterial plasmids. The majority of tested bifidobacteria do not appear to harbour plasmids (less than 20% of the strains analysed have been shown to harbour such extrachromosomally replicating entities) (Sgorbati *et al.*, 1982), and it is probably for this reason that there are fewer plasmids sequenced as compared to the number of sequenced genomes (See Chapter I, Table 1.2).

Bifidobacterial plasmid sequence analysis and subsequent functional characterisation of their corresponding replication functions has allowed the construction of novel genetic tools for use in this group of gut commensals (Hirayama *et al.*, 2012; O'Connell Motherway *et al.*, 2009; Ruiz *et al.*, 2013; Alvarez-Martin *et al*., 2008; Cronin *et al*., 2007). Such tools, when combined with bifidobacterial genome sequencing projects, facilitate functional genomics approaches that are central to further our understanding on how these bacteria colonise and persist in the highly competitive environment of the gastrointestinal tract (GIT).

Glycosyl hydrolases (GHs) are essential for bifidobacteria in order to enzymatically access complex diet- or host-derived carbohydrates that are present in the GIT environment. Tables 1.4-1.8 in Chapter I presents a compilation of predicted

bifidobacterial GHs as based on Cazy website information (which hosts information on carbohydrate active enzymes (http://www.cazy.org/)). Compared to a previously published compilation (Pokusaeva *et al.*, 2009), a significant increase is notable in the number of putative GHs, which is not surprising as it corresponds to the rise in completely sequenced bifidobacterial genomes. It is important to understand how these GHs function in bifidobacteria so that novel bifidogenic prebiotics can be discovered. The first step in discovering such prebiotics is to elucidate the saccharidic substrates associated with the identified GH-encoding genes, and to discover the metabolic functions and regulation of these genes. Chapters III and IV of this thesis outline research on the function and regulation of two gene clusters involved in the metabolism of raffinose-like sugars and melezitose. The obtained data can be used to select potential prebiotics, which will still require further *in vivo* evaluation with regards to their efficacy and selectivity as bifidogenic compounds. In addition, the acquired information on the enzymatic activity and specificity of the characterised GHs may allow the production of novel prebiotics through transglycosylation. Transglycosylation is the ability of so-called retaining GHs to catalyse carbohydrate condensation reactions by the double displacement mechanism with retention of configuration at the anomeric centre to yield elongated oligosaccharides with new glycosidic linkages (Withers, 2001).

In Chapter II of this thesis, the ability of bifidobacteria to produce conjugated linoleic acid (CLA) from linoleic acid (LA; a dienoic unsaturated fatty acid) was investigated so as to characterise the enzymatic activity which allows certain bifidobacterial strains to produce this compound (O'Connell *et al.*, 2013a; Rosberg-Cody *et al.*, 2004). CLA is claimed to promote to host health, for example due to its perceived action in anti-obesity (Chen *et al.*, 2012), its anti-carcinogenesis activity

(Pierre *et al.*, 2013), its immunomodulatory role (Bassaganya-Riera *et al.*, 2012) and its involvement in bone formation (Park *et al.*, 2013).

The precise mechanism by which CLA is produced by certain bacteria is as yet unknown, although several hypotheses have been put forward to explain this phenomenon. For example, in *Lactobacillus acidophilus* AKU 1137 it is speculated that hydroxy fatty acid formation from linoleic acid is the first step in CLA formation (Ogawa *et al.*, 2001). These authors found that when hydroxy fatty acids were introduced to washed *Lb. acidophilus* AKU 1137 cells, these hydroxy fatty acids are transformed to CLA isomers via an LA isomerase (LAI).

It has been shown previously that myosin cross reactive antigen (MCRA)-like proteins from various microorganisms, including bifidobacteria, show over 50 % sequence similarity to the primary sequence of LAI proteins from *Lb. acidophilus* and *Lactobacillus reuteri* PYR8 (Rosson *et al.*, 2001). These latter enzymes are believed to catalyse the conversion of LA to *cis*-9 *trans*-11 (c9,t11) CLA. However, in Chapter II of this thesis it is shown that the MCRA-encoding gene of *B. breve* NCFB 2258 is not responsible for CLA production, since an insertion mutation in the MCRA-encoding gene, does not affect the CLA-producing abilities of this *B. breve* strain. Instead, we found that the MCRA-encoded gene product functions as an oleate hydratase, which is involved in the production of 10-hydroxy fatty acids from oleic acid (O'Connell *et al.*, 2013a). This finding supports previously reported data (Rosberg-Cody *et al.*, 2011), in which it was shown that the MCRA protein encoded by *B. breve* NCFB 2258, when heterologously expressed in *Lactococcus lactis* and *Corynebacterium glutamicum*, possesses hydratase activity (catalysing the production of 10-hydroxy fatty acids from oleic acid). Despite suggestive data (Ogawa *et al.*, 2001), which would imply the presumed bifidobacterial MCRA

protein and its associated linoleic acid isomerase activity in the formation of CLA from LA, our findings show that this enzyme is not involved in CLA production. Therefore, the actual metabolic pathway of CLA production in bifidobacteria is still elusive.

Previous findings (Rosberg-Cody *et al.*, 2011), which were corroborated by our data (Chapter II of this thesis), have shown that the MCRA protein of *B. breve* NCFB 2258 is involved in the provision of increased tolerance to particular solvents, such as ethanol. Ethanol is a short chain alcohol and ethanol exposure results in changes in membrane fatty acid composition, structure and membrane fluidity in bacteria (Liu & Qureshi, 2009). In a previous study deletion of the MCRA-encoding gene in *Lactobacillus acidophilus* NCFM resulted in a strain that exhibited reduced growth in the presence of lactate, acetate and salt (O'Flaherty & Klaenhammer, 2010). The mechanisms by which Gram positive bacteria tolerate organic solvents have yet to be elucidated (Torres *et al.*, 2011), but tolerance may be provided by the induction of a general stress regulon (Sardessai & Bhosle, 2002), production of organic solvent-emulsifying or deactivating enzymes (Moriya *et al.*, 1995) and/or an active solvent efflux pump (Inoue & Horikoshi, 1991; Moriya *et al.*, 1995). Based on a previous study (Rosberg-Cody *et al.*, 2011) and our own findings it may thus be concluded that the bifidobacterial MCRA-like protein is involved in the adjustment of the fatty acid composition of the lipid bi-layer in order to provide increased tolerance to solvents such as ethanol.

Due to the reported health-promoting properties of CLA it is desirable to develop functional foods which contain high levels of these bioactive compounds in order to elicit a beneficial effect on human health. LA is naturally present in relatively low amounts in milk and dairy products, meat and meat products of ruminant animals

(Bhattacharya *et al.*, 2006). The average dose of CLA reached per day in the human diet is 650 mg (Jutzeler van Wijlen & Colombani, 2010) and this level is not high enough to confer any beneficial effects on the host. For example, in order to achieve beneficial effects in terms of body fat reduction, daily doses of 3-4 g are advised (Bhattacharya *et al.*, 2006). Ways of increasing CLA in meat and dairy products may be possible by means of addition of oilseeds, vegetable or fish oil to animal feed (Duckett & Gillis, 2010). The ability of certain gut commensals to produce CLA presents another route by which CLA demands can be met and thus can be considered an important probiotic characteristic of such bacteria (O'Shea *et al.*, 2012).

Several bifidobacterial strains have been associated with probiotic activities, and have been used as live microbial food ingredients that, when ingested, have a beneficial effect on human health (Salminen *et al.*, 1998). There are many health benefits associated with bifidobacteria, such as inhibition of pathogens (Fanning *et al.*, 2012; Fukuda *et al.*, 2011), alleviating lactose intolerance (Almeida *et al.*, 2012) and reducing serum cholesterol (Bordoni *et al.*, 2013). The health-promoting activities of probiotic bacteria may be stimulated by prebiotics, which for this reason are defined as non-digestible (by the human host) food ingredients that beneficially affect the host by selectively stimulating the growth and/or activity of one or a limited number of probiotic bacteria in the colon (Gibson & Roberfroid, 1995). It has been shown that microbiota changes with age, with bifidobacteria being predominant during the stage when (breast) milk is the main food source, while their relative numbers will decline following weaning (Yatsunenko *et al.*, 2012). Microbiota composition has been shown to influence general health and it has been shown that particular alterations in intestinal microbiota composition are associated with several

chronic conditions, including obesity and inflammatory diseases in the elderly (Claesson *et al.*, 2012). This study has linked diet, microbiota and health status, and therefore has highlighted a specific and very promising role for prebiotics in modulating the microbiota.

In order to gain an insight into how potential prebiotic compounds affect host microbiota composition and subsequently host health, we must first investigate how these prebiotics are metabolised by the targeted probiotic bacteria. Chapter III of this thesis describes the identification and characterisation of two gene clusters responsible for the metabolism of the saccharidic compounds raffinose and raffinose-like carbohydrates, and melezitose, all of which have prebiotic potential, in *B. breve* UCC2003. Stachyose and raffinose are present in a wide variety of plants (French, 1954), while the related sugar melibiose (though not a member of the raffinose family) is also found in many plants and is particularly abundant in soybean roots and stems (Rehms & Barz, 1995). Melezitose is found in honeydew and manna, which are sugar-rich liquid and solid deposits, respectively, associated with leaves and branches of various trees and shrubs (Bacon & Dickinson, 1957).

In this study we identified gene clusters involved in the uptake and catabolism of raffinose-containing carbohydrates and melezitose in *B. breve* UCC2003. *In silico* analyses, insertion mutagenesis, protein purification and enzyme kinetics allowed a detailed characterisation of the functionality of various genes and corresponding protein products of these clusters (O'Connell *et al.*, 2013b). An α-galactosidase (encoded by *rafA*) and a putative solute binding protein (encoded by *rafB*) were shown to be specified by the raffinose-utilisation gene cluster. The purified RafA was shown to hydrolyse stachyose, raffinose and melibiose to produce sucrose and

galactose, while also cleaving melibiose to its monosaccharide constituents glucose and galactose. Kinetic studies identified that the preferred bond cleaved by RafA was melibiose, followed by raffinose (O'Connell *et al.*, 2013b). Our findings indicate that many bifidobacteria metabolise raffinose-like sugars by means of a metabolic route, which is different to that known for *E. coli* (Aslanidis & Schmitt, 1990). Raffinose metabolism has previously been investigated in *Lactobacillus plantarum* ATCC 8014 (Silvestroni *et al.*, 2002), and from these and our data it is evident that the raffinose utilisation cluster of *Lb. plantarum* differs substantially from that found in bifidobacteria.

In the case of the melezitose utilisation cluster, an α-galactosidase (specified by *melE*) and α-glucosidase (specified by *melD*) were shown to be specified by the melezitose utilisation gene cluster. Purified recombinant MelD protein was used to demonstrate that this protein has both α-(1→2) and α-(1→3) glucosyl hydrolase activities. MelE was shown not to exhibit hydrolytic activity towards raffinose, stachyose or melibiose, yet this protein as well as RafA were demonstrated to exhibit hydrolytic activity towards synthetic α-(1→4)- and α-(1→3)-galactobiose. Kinetic studies identified that the preferred bond cleaved by MelD is the α-(1→2)-glucosidic linkage present in sucrose (O'Connell *et al.*, 2013b). In terms of melezitose metabolism in other bacteria there is a paucity of information. Melezitose metabolism has only been investigated in the yeast *Saccharomyces cerevisiae* (Hwang & Lindegren, 1964), revealing a melezitose metabolising system employing an α-glucosidase or melezitase.

Other carbohydrate metabolising systems have been studied in *B. breve* UCC2003 (Maze *et al.*, 2007; O'Connell *et al.*, 2013b; O'Connell Motherway *et al.*, 2008; O'Connell Motherway *et al.*, 2011; Pokusaeva *et al.*, 2010; Pokusaeva *et al.*, 2011;

Ryan *et al.*, 2005). The study presented in Chapter III of this thesis thus provides relevant information as to how bifidobacteria metabolise raffinose-containing carbohydrates and melezitose. These pathways dedicated to the uptake and metabolism of potential prebiotics may represent common elements in the diet of the host of such bacteria, while it may also allow certain species/strains to colonise the gut or obtain higher numbers in the presence of these carbohydrates (van Zanten *et al.*, 2012).

It has to be realised that the study presented here is based on *in vitro* assays and were intended to gain more insight into the prebiotic potential of these carbohydrates. The obtained information are therefore a prelude to *in vivo* trials using animal models that may be followed by human clinical trials. Such studies will provide more realistic information regarding food ingredients, perhaps in combination with selective probiotics, to be incorporated into our diet in order to influence microbiota and, ultimately, host health.

The fourth chapter of this thesis is a follow on from Chapter III and describes investigations that were aimed to achieve a better understanding of how *B. breve* UCC2003 regulates the transcription of the gene clusters involved in raffinose and melezitose metabolism. From the observed findings we deduce that the gene cluster associated with melezitose metabolism is negatively regulated by a LacI-type transcriptional regulator designated MelR1. A second LacI-type regulator, MelR2, which is encoded by a gene adjacent to the *melR1* gene, was shown to regulate two genes thought to be responsible for the uptake of a melezitose-like carbohydrate which contains an α-galactose moiety. We were unable to test this theory as such carbohydrates are not commercially available. Furthermore, a ROK-type protein,

RafR, was shown to act as a transcriptional activator of the *rafBCD* cluster, the first such activity described for bifidobacteria. A ROK-type protein has previously been linked to positive transcriptional control in *Corynebacterium glutamicum* 13032 (Ruckert *et al.*, 2008), since a deletion in the gene encoding the ROK-type protein, CysR, rendered this strain incapable of utilising sulphate or sulphonates as sole source of sulphur.

In Chapter IV we also report that melezitose is preferentially utilised over raffinose, indicating a form of carbohydrate utilisation hierarchy in bifidobacteria. Specific elements of the PEP-PTS system are in many bacteria responsible for carbon catabolite control (Deutscher, 2008), however, this does not seem to be case in bifidobacteria as regulatory proteins which are essential for this system appear to be absent. Therefore, the manner by which catabolite repression is achieved by bifidobacteria is yet to be elucidated. Transcription of bifidobacterial gene clusters involved in carbohydrate metabolism seems to be tightly regulated by dedicated regulators, in many cases LacI-type repressors, yet several observations have been made which suggest that these clusters are also subject to a higher order control system that possibly enforces global regulation (Parche *et al.*, 2006). Further studies will need to be done in order to elucidate this system.

Finally, Chapter V describes the characterisation of the replication functions associated with a novel bifidobacterial megaplasmid and attempts to create an *E. coli*/bifidobacterial shuttle vector as a novel genetic tool to genetically manipulate bifidobacteria. Bifidobacteria are known to be genetically recalcitrant, while the arsenal of genetic tools, such as cloning, conditionally replicating and (inducible) expression vectors, is limited or non-existent. Various studies have been dedicated to overcome the barrier imposed by restriction/modification systems in achieving

higher levels of transformation (Hirayama *et al.*, 2012; O'Connell Motherway *et al.*, 2009; Ruiz *et al.*, 2013). In addition, quite a number of papers have been published on the identification of bifidobacterial plasmids and the construction of vectors that can shuttle between *E. coli* and particular bifidobacterial species/strains (Alvarez-Martin *et al.*, 2008; Cronin *et al.*, 2007; Klijn *et al.*, 2006; Lee & O'Sullivan, 2006; Matsumura *et al.*, 1997; Missich *et al.*, 1994; Rossi *et al.*, 1996; Rossi *et al.*, 1998; Sangrador-Vegas *et al.*, 2007; Shkoporov *et al.*, 2008).

In the latter context, the discovery of a megaplasmid in *B. breve* JCM 7017 provided opportunities to expand the currently available genetic tool box for bifidobacteria with vectors that would allow the cloning and stable maintenance of large DNA fragments. In order to study the replication functions of the megaplasmid, *in silico* analyses were performed which identified two genes, $repA_{7017}$ and $repB_{7017}$, encoding putative replication proteins. The $RepA_{7017}$ protein was shown to bind the upstream region of its encoding gene, which was shown to contain a series of direct repeats as well as its promoter. Attempts to utilise this information for the construction of a bifidobacterial/*E. coli* shuttle vector unfortunately failed. This could be due to the fact that not all replication functions were present in our constructs or that these replication functions were present but were dysfunctional.

The work presented in this thesis has provided novel information not only on how bifidobacteria utilise various carbohydrates with potential prebiotic activity, but also how they regulate the genes necessary for the metabolism of these saccharides, while also showing promising results for the production of future genetic tools for the genetic manipulation of bifidobacteria. In terms of CLA production our research has provided some further insights into fatty acid metabolism, however, the precise mechanism of CLA production by bifidobacteria is still a mystery.

6.0 REFERENCES

Almeida, C. C., Lorena, S. L. S., Pavan, C. R., Akasaka, H. M. I. & Mesquita, M. A. (2012). Beneficial Effects of Long-Term Consumption of a Probiotic Combination of *Lactobacillus casei Shirota* and *Bifidobacterium breve* Yakult May Persist After Suspension of Therapy in Lactose-Intolerant Patients. *Nutrition in Clinical Practice* **27**, 247-251.

Alvarez-Martin, P., Belen Florez, A., Margolles, A., del Solar, G. & Mayo, B. (2008). Improved cloning vectors for bifidobacteria, based on the *Bifidobacterium catenulatum* pBC1 replicon. *Applied and Environmental Microbiology* **74**, 4656-4665.

Aslanidis, C. & Schmitt, R. (1990). Regulatory elements of the raffinose operon: nucleotide sequences of operator and repressor genes. *Journal of Bacteriology* **172**, 2178-2180.

Bacon, J. S. & Dickinson, B. (1957). The origin of melezitose: a biochemical relationship between the lime tree (*Tilia* spp.) and an aphis (*Eucallipterus tiliae L.*). *The Biochemical Journal* **66**, 289-297.

Bassaganya-Riera, J., Hontecillas, R., Horne, W. T., Sandridge, M., Herfarth, H. H., Bloomfeld, R. & Isaacs, K. L. (2012). Conjugated linoleic acid modulates

immune responses in patients with mild to moderately active Crohn's disease. *Clinical Nutrition* **31**, 721-727.

Bhattacharya, A., Banu, J., Rahman, M., Causey, J. & Fernandes, G. (2006). Biological effects of conjugated linoleic acids in health and disease. *The Journal of Nutritional Biochemistry* **17**, 789-810.

Bordoni, A., Amaretti, A., Leonardi, A., Boschetti, E., Danesi, F., Matteuzzi, D., Roncaglia, L., Raimondi, S. & Rossi, M. (2013). Cholesterol-lowering probiotics: *in vitro* selection and *in vivo* testing of bifidobacteria. *Applied Microbiology and Biotechnology* **97**, 8273-8281.

Chen, S.-C., Lin, Y.-H., Huang, H.-P., Hsu, W.-L., Houng, J.-Y. & Huang, C.-K. (2012). Effect of conjugated linoleic acid supplementation on weight loss and body fat composition in a Chinese population. *Nutrition* **28**, 559-565.

Claesson, M. J., Jeffery, I. B., Conde, S., Power, S.E., O' Connor, E.M., Cusack, S., Harris, H.M., Coakley, M., Lakshminarayanan, B., O' Sullivan, O., Fitzgerald, G.F., Deane, J., O' Connor, M., Harnedy, N., O' Connor, K., O' Mahony, D., van Sinderen, D., Wallace, M., Brennan, L., Stanton, C., Marchesi, J.R., Fitzgerald, A.P., Shanahan, F., Hill, C., Ross, R.P. & O' Toole P.W. (2012). Gut microbiota composition correlates with diet and health in the elderly. *Nature* **488**, 178-184.

Cronin, M., Knobel, M., O'Connell Motherway, M., Fitzgerald, G. F. & van Sinderen, D. (2007). Molecular Dissection of a Bifidobacterial Replicon. *Applied and Environmental Microbiology* **73**, 7858-7866.

Duckett, S. K. & Gillis, M. H. (2010). Effects of oil source and fish oil addition on ruminal biohydrogenation of fatty acids and conjugated linoleic acid formation in beef steers fed finishing diets. *Journal of Animal Science* **88**, 2684-2691.

Fanning, S., Hall, L. J., Cronin, M. Zomer, A., MacSharry, J., Goulding, D., O'Connell Motherway, M., Shanahan, F., Nally, K., Dougan, G. & van Sinderen, D. (2012). Bifidobacterial surface-exopolysaccharide facilitates commensal-host interaction through immune modulation and pathogen protection. *Proceedings of the National Academy of Sciences* **109**, 2108-2113.

French, D. (1954). The Raffinose Family of Oligosaccharides. In *Advances in Carbohydrate Chemistry*, pp. 149-184. Edited by L. W. Melville: Academic Press.

Fukuda, S., Toh, H., Hase, K. Fukuda, S., Toh, H., Hase, K., Oshima, K., Nakanishi, Y., Yoshimura, K.,Tobe, T.,Clarke,J.M., Topping, D.L.,Suzuki, T., Taylor, T.D., Itoh, K., Kikuchi, J., Morita, H.,Hattori, M. & Ohno, H. (2011). Bifidobacteria can protect from enteropathogenic infection through production of acetate. *Nature* **469**, 543-547.

Gibson, G. R. & Roberfroid, M. B. (1995). Dietary modulation of the human colonic microbiota: introducing the concept of prebiotics. *The Journal of nutrition* **125**, 1401-1412.

Hirayama, Y., Sakanaka, M., Fukuma, H., Murayama, H., Kano, Y., Fukiya, S. & Yokota, A. (2012). Development of a double-crossover markerless gene deletion system in *Bifidobacterium longum*: functional analysis of the alpha-galactosidase gene for raffinose assimilation. *Applied and Environmental Microbiology* **78**, 4984-4994.

Hwang, D. S. & Lindegren, C. C. (1964). Palatinose Element of the Receptor of the Melezitose Locus in Saccharomyces. *Nature* **203**, 791-792.

Inoue, A. & Horikoshi, K. (1991). Estimation of solvent-tolerance of bacteria by the solvent parameter log P. *Journal of Fermentation and Bioengineering* **71**, 194-196.

Jutzeler van Wijlen, R. P. & Colombani, P. C. (2010). Grass-based ruminant production methods and human bioconversion of vaccenic acid with estimations of maximal dietary intake of conjugated linoleic acids. *International Dairy Journal* **20**, 433-448.

Klijn, A., Moine, D., Delley, M., Mercenier, A., Arigoni, F. & Pridmore, R. D. (2006). Construction of a Reporter Vector for the Analysis of *Bifidobacterium longum* Promoters. *Applied and Environmental Microbiology* **72**, 7401-7405.

Lee, J. H. & O'Sullivan, D. J. (2006). Sequence analysis of two cryptic plasmids from *Bifidobacterium longum* DJO10A and construction of a shuttle cloning vector. *Applied and Environmental Microbiology* **72**, 527-535.

Liu, S. & Qureshi, N. (2009). How microbes tolerate ethanol and butanol. *New Biotechnology* **26**, 117-121.

Matsumura, H., Takeuchi, A. & Kano, Y. (1997). Construction of *Escherichia coli-Bifidobacterium longum* shuttle vector transforming *B. longum* 105-A and 108-A. *Bioscience, Biotechnology, and Biochemistry* **61**, 1211-1212.

Maze, A., O'Connell Motherway, M., Fitzgerald, G. F., Deutscher, J. & van Sinderen, D. (2007). Identification and characterisation of a fructose phosphotransferase system in *Bifidobacterium breve* UCC2003. *Applied and Environmental Microbiology* **73**, 545-553.

Missich, R., Sgorbati, B. & LeBlanc, D. J. (1994). Transformation of *Bifidobacterium longum* with pRM2, a constructed *Escherichia coli-B. longum* shuttle vector. *Plasmid* **32**, 208-211.

Moriya, K., Yanigitani, S., Usami, R. & Horikoshi, K. (1995). Isolation and some properties of an organic solvent marine bacterium degrading cholesterol. *Journal of Marine Biotechnology* **2**, 131-133.

O'Connell, K. J., O'Connell Motherway, M., Hennessey, A. A., Brodhun, F., Ross, R. P., Feussner, I., Stanton, C., Fitzgerald, G. F. & van Sinderen, D. (2013a). Identification and characterisation of an oleate hydratase-encoding gene from *Bifidobacterium breve*. *Bioengineered* **4**, 313-321.

O'Connell, K. J., O'Connell Motherway, M., O'Callaghan, J., Fitzgerald, G. F., Ross, R. P., Ventura, M., Stanton, C. & van Sinderen, D. (2013b). Metabolism of four α-glycosidic linkage-containing oligosaccharides by *Bifidobacterium breve* UCC2003. *Applied and Environmental Microbiology* **79**, 6280-6292.

O'Connell Motherway, M., Fitzgerald, G. F., Neirynck, S., Ryan, S., Steidler, L. & van Sinderen, D. (2008). Characterisation of ApuB, an extracellular type II amylopullulanase from *Bifidobacterium breve* UCC2003. *Applied and Environmental Microbiology* **74**, 6271-6279.

O'Connell Motherway, M., O'Driscoll, J., Fitzgerald, G. F. & van Sinderen, D. (2009). Overcoming the restriction barrier to plasmid transformation and targeted mutagenesis in *Bifidobacterium breve* UCC2003. *Microbial Biotechnology* **2**, 321-332.

O'Connell Motherway, M., Fitzgerald, G. F. & van Sinderen, D. (2011). Metabolism of a plant derived galactose-containing polysaccharide by *Bifidobacterium breve* UCC2003. *Microbial Biotechnology* **4**, 403-416.

O'Flaherty, S. J. & Klaenhammer, T. R. (2010). Functional and phenotypic characterisation of a protein from *Lactobacillus acidophilus* involved in cell morphology, stress tolerance and adherence to intestinal cells. *Microbiology* **156**, 3360-3367.

O'Shea, E. F., Cotter, P. D., Stanton, C., Ross, R. P. & Hill, C. (2012). Production of bioactive substances by intestinal bacteria as a basis for explaining probiotic mechanisms: Bacteriocins and conjugated linoleic acid. *International Journal of Food Microbiology* **152**, 189-205.

Ogawa, J., Matsumura, K., Kishino, S., Omura, Y. & Shimizu, S. (2001). Conjugated linoleic acid accumulation via 10-hydroxy-12-octadecaenoic acid during microaerobic transformation of linoleic acid by *Lactobacillus acidophilus*. *Applied and Environmental Microbiology* **67**, 1246-1252.

Park, Y., Kim, J., Scrimgeour, A. G., Condlin, M. L., Kim, D. & Park, Y. (2013). Conjugated linoleic acid and calcium co-supplementation improves bone health in ovariectomised mice. *Food Chemistry* **140**, 280-288.

Pierre, A.-S., Minville-Walz, M., Fèvre, C. Hichami, A., Gresti, J., Pichon, L., Bellenger, S., Bellenger, J., Ghiringhelli, F., Narce, M. & Rialland, M. (2013). Trans-10, cis-12 conjugated linoleic acid induced cell death in human colon cancer cells through reactive oxygen species-mediated ER stress. *Molecular and Cell Biology of Lipids* **1831**, 759-768.

Pokusaeva, K., O'Connell Motherway, M., Zomer, A., Fitzgerald, G. F. & van Sinderen, D. (2009). Characterisation of two novel alpha-glucosidases from *Bifidobacterium breve* UCC2003. *Applied and Environmental Microbiology* **75**, 1135-1143.

Pokusaeva, K., Neves, A. R., Zomer, A., O'Connell Motherway, M., MacSharry, J., Curley, P., Fitzgerald, G. F. & van Sinderen, D. (2010). Ribose utilisation by the human commensal *Bifidobacterium breve* UCC2003. *Microbial Biotechnology* **3**, 311-323.

Pokusaeva, K., O'Connell Motherway, M., Zomer, A., Macsharry, J., Fitzgerald, G. F. & van Sinderen, D. (2011). Cellodextrin utilisation by *Bifidobacterium breve* UCC2003. *Applied and Environmental Microbiology* **77**, 1681-1690.

Rehms, H. & Barz, W. (1995). Degradation of stachyose, raffinose, melibiose and sucrose by different tempe-producing *Rhizopus fungi*. *Applied Microbiology and Biotechnology* **44**, 47-52.

Rosberg-Cody, E., Ross, R. P., Hussey, S., Ryan, C. A., Murphy, B. P., Fitzgerald, G. F., Devery, R. & Stanton, C. (2004). Mining the microbiota of the neonatal gastrointestinal tract for conjugated linoleic acid-producing bifidobacteria. *Applied and Environmental Microbiology* **70**, 4635-4641.

Rosberg-Cody, E., Liavonchanka, A., Gobel, C., Ross, R. P., O'Sullivan, O., Fitzgerald, G. F., Feussner, I. & Stanton, C. (2011). Myosin-cross-reactive antigen (MCRA) protein from *Bifidobacterium breve* is a FAD-dependent fatty acid hydratase which has a function in stress protection. *BMC Biochemistry* **12**, 9.

Rossi, M., Brigidi, P., Gonzalez Vara y Rodriguez, A. & Matteuzzi, D. (1996). Characterisation of the plasmid pMB1 from *Bifidobacterium longum* and its use for shuttle vector construction. *Research in Microbiology* **147**, 133-143.

Rossi, M., Brigidi, P. & Matteuzzi, D. (1998). Improved cloning vectors for *Bifidobacterium* spp. *Letters in Applied Microbiology* **26**, 101-104.

Rosson, R. A., Grund, A., Deng, M. & Sanchez-Riera, F. (2001). Linoleate isomerase. *World Patent* **100846**, 30.

Ruckert, C., Milse, J., Albersmeier, A., Koch, D., Puhler, A. & Kalinowski, J. (2008). The dual transcriptional regulator CysR in *Corynebacterium glutamicum* ATCC 13032 controls a subset of genes of the McbR regulon in response to the availability of sulphide acceptor molecules. *BMC Genomics* **9**, 483.

Ruiz, L., O'Connell Motherway, M., Lanigan, N. & van Sinderen, D. (2013). Transposon mutagenesis in *Bifidobacterium breve*: construction and characterisation of a Tn5 transposon mutant library for *Bifidobacterium breve* UCC2003. *PLoS One* **8**, e64699.

Ryan, S. M., Fitzgerald, G. F. & van Sinderen, D. (2005). Transcriptional Regulation and Characterisation of a Novel β-Fructofuranosidase-Encoding Gene from *Bifidobacterium breve* UCC2003. *Applied and Environmental Microbiology* **71**, 3475-3482.

Salminen, S., Bouley, C., Boutron-Ruault, M.C., Cummings, J.H., Franck, A., Gibson, G.R., Isolauri, E., Moreau, M.C., Roberfroid, M. & Rowland, I. (1998). Functional food science and gastrointestinal physiology and function. *The British Journal of Nutrition* **80**, 147-171.

Sangrador-Vegas, A., Stanton, C., van Sinderen, D., Fitzgerald, G. F. & Ross, R. P. (2007). Characterisation of plasmid pASV479 from *Bifidobacterium*

pseudolongum subsp. *globosum* and its use for expression vector construction. *Plasmid* **58**, 140-147.

Sardessai, Y. & Bhosle, S. (2002). Tolerance of bacteria to organic solvents. *Research in Microbiology* **153**, 263-268.

Sgorbati, B., Scardovi, V. & Leblanc, D. J. (1982). Plasmids in the Genus *Bifidobacterium*. *Journal of General Microbiology* **128**, 2121-2131.

Shkoporov, A. N., Efimov, B. A., Khokhlova, E. V., Steele, J. L., Kafarskaia, L. I. & Smeianov, V. V. (2008). Characterisation of plasmids from human infant *Bifidobacterium* strains: sequence analysis and construction of *E. coli*-*Bifidobacterium* shuttle vectors. *Plasmid* **60**, 136-148.

Silvestroni, A., Connes, C., Sesma, F., De Giori, G. S. & Piard, J. C. (2002). Characterisation of the melA locus for alpha-galactosidase in *Lactobacillus plantarum*. *Applied and Environmental Microbiology* **68**, 5464-5471.

Torres, S., Pandey, A. & Castro, G. R. (2011). Organic solvent adaptation of Gram positive bacteria: applications and biotechnological potentials. *Biotechnology Advances* **29**, 442-452.

van Zanten, G. C., Knudsen, A., Roytio, H., Forssten, S., Lawther, M., Blennow, A., Lahtinen, S.J., Jakobsen, M., Svensson, B. & Jespersen, L. (2012). The effect of selected synbiotics on microbial composition and short-chain fatty acid production in a model system of the human colon. *PLoS One* **7**, e47212.

Ventura, M., Turroni, F., Lugli, G. A. & van Sinderen, D. (2014). Bifidobacteria and humans: our special friends, from ecological to genomics perspectives. *Journal of the Science of Food and Agriculture* **94**, 163-168.

Withers, S. G. (2001). Mechanisms of glycosyl transferases and hydrolases. *Carbohydrate Polymers* **44**, 325-337.

Yatsunenko, T., Rey, F.E., Manary, M.J., Trehan I, Dominguez-Bello, M.G., Contreras, M., Magris, M., Hidalgo, G., Baldassano, R.N., Anokhin, A.P., Heath, A.C., Warner, B., Reeder, J., Kuczynski, J., Caporaso, J.G., Lozupone, C.A., Lauber, C., Clemente, J.C., Knights, D., Knight, R. & Gordon, J.I. (2012). Human gut microbiome viewed across age and geography. *Nature* **486**, 222-227.

ACKNOWLEDGEMENTS

There are many people I would like to acknowledge for their contribution throughout the PhD process,

My supervisor Professor Douwe van Sinderen whose words of wisdom, encouragement, patience and belief in me was pivotal to my success, I could not have worked in a better lab and I could not have wished for a better supervisor.

Everyone in lab 4.25 who over the years have become firm friends, thank you for supporting me in stressful times and celebrating with me when things went right, especially Dr. Mary O' Connell Motherway for all your help from queries about my lab work, to writing, I really appreciated your guidance, it was invaluable to me.

My co-supervisor Professor Catherine Stanton, for supervision and helpful discussions, especially regarding the CLA part of my studies.

Dr. Paul O' Toole who was on my thesis committee, read and corrected my reports and came to my poster sessions.

The staff in the Microbiology Department for help with strains, computer issues, orders etc. Paddy O' Reilly, Dan Walsh, Collette O' Gorman, Margaret Higgins, Pat Higgins, Hilda Bohane, Liam Burgess, to name but a few.

I would like to acknowledge the financial contributions of the TOMAR trust and Science Foundation Ireland through the Alimentary Pharmabiotic Centre.

Everyone I collaborated with in various parts of my chapters Dr. Alan Hennessey who had an input in my CLA analysis, Professor Ivo Feussner and Florian Brodhun who performed hydratase analysis, and contributed to my very first paper which I

was so proud to achieve, Dr. John O' Callaghan for help with kinetics, Dr. Stephen Cunningham, Dr. Marian Keane and Prof. Lokesh Joshi of the Alimentary Glycoscience Research Cluster, NUIG, Ireland, for kindly supplying carbohydrates and last but not least Francesca Bottacini for helping me with my bioinformatics.

My parents Michael and Siobhan O' Connell, who always told me I could do anything I set my mind to, have always loved, supported and believed in me. I couldn't have done any of this without your belief and encouragement, it was key to my success.

My three sisters Kim, Vikki and Eileen O' Connell who are always there to spur me on, I am very lucky to have sisters like you.

My nephew and godson Kieran who never fails to make me smile.

My fiancé Chris Martin who has supported me every step of the way from my school days to University and much more. You are my rock, you bring out the best in me.

I would also like to thank Chris' family for their support and interest in my work through the years.

My Uncle Denis Slattery who is such an inspirational person and somebody I aspire to be like.

My Uncle Jimmy Slattery…Great minds think alike.

My Cousin Adam Slattery, a fellow scientist, who has always encouraged my love of science and also helped me with Physics during my degree.

My dearest friends Kathy Barriscale, Martina Blake, Deirdre Crowley and Claire Creedon, I appreciate having such great friends.